

**République Algérienne Démocratique et Populaire Ministère de
l'Enseignement Supérieur et de la Recherche Scientifique**

UNIVERSITE SAAD DAHLEB DE BLIDA

Faculté des Sciences

Département d'Informatique



MEMOIRE DE FIN D'ETUDES

Pour l'obtention du diplôme de Master en Informatique

Option : Ingénierie des Logiciels

Date de soutenance : 03 Octobre 2021

THEME

*Evolution des intérêts de l'utilisateur dans son réseau
social*

**Organisme d'accueil : Centre de Recherche sur l'Information Scientifique et
Technique (CERIST)**

Réalisé par :

BRINIS IHCENE et RAMDANI DALILA

Promotrice : Mme BENBLIDIA NADJIA (Université de BLIDA)

Encadreur : Mme BOULKRINAT NOUR EL HOUDA (CERIST)

Président : Mme OUKID SALIHA (Université de BLIDA)

Examineur : Mme LAHIANI NESRINE (Université de BLIDA)

Année Universitaire 2020 / 2021

Remerciements

Avant tout développement sur cette expérience, il apparaît opportun de commencer par des remerciements, à ceux qui nous ont beaucoup appris au cours de cette année, et même à ceux qui ont eu la gentillesse de faire de cette expérience un moment très profitable.

Nous tenons tout d'abord à remercier Dieu le tout puissant, qui nous a donné la force et la patience d'accomplir ce travail dans les meilleures conditions.

Nos plus sincères remerciements, à notre encadreur Mme Boulkrinat Nour El Houda pour son attention, son orientation, son aide pendant la réalisation de ce travail et pour être source d'information et de communication sans hésiter à aucun moment de consacrer une part de son temps précieux.

Nos remerciant très chaleureusement aussi, Mme Benblidia Nadja, notre promotrice pour sa confiance et ses encouragements.

Nous remercions tous les membres de jury avoir accepté d'examiner et de juger ce travail.

Finalement nous remercions nos parents, et nos proches pour leur soutien moral et matériel.

Dédicaces

J'ai le grand plaisir de dédié ce modeste travail :

A ma très chère mère, qui me donne toujours l'espoir de vivre et qui n'a jamais cessé de prier pour moi.

A mon très cher père, pour ses encouragements, son soutien, surtout pour son amour et son sacrifice afin que rien n'entrave le déroulement de mes études.

A mon cher frère Lyes et ma chère sœur Rania.

A mes grands-mères Zineb, Yamina et à mes grands-pères Mohamed, Mahfoud. Que Dieu leur donne une longue et joyeuse vie.

A tous membre de ma famille.

A mon encadreur Mme Boulkrinat Nour El Houda, qui m'a donnée l'effort et le soutien moral pour la réalisation de ce projet.

A mon promotrice Mme Benblidia Nadjia.

A mon binôme et ma très chère amie Dalila, pour son soutien moral, sa patience et sa compréhension tout au long de ce projet.

A tous mes amis de promotion de 2^{ème} année Master en Ingénierie des Logiciels.

A mes meilleures amies proches : Ghania, Lamia et Amel.

A mes très chers oiseaux au mon cœur.

Ihcene

Dédicaces

J'ai le grand plaisir de dédié ce modeste travail :

A ma mère qui m'a encouragé durant toutes mes études, et qui sans elle, ma réussite n'aura pas eu lieu. Qu'elle trouve ici mon amour et mon affection,

A mon père qui est toujours disponible pour nous, et prêt à nous aider, afin que rien n'entrave le déroulement de mes études,

Je lui confirme mon attachement et mon profond respect,

A ma sœur chérie Siham pour m'avoir conseillé, encouragé, soutenu au long de mon parcours d'étude,

A ma grande mère la personne plus importante de ma vie qui m'a soutenu,

A ma tante Karima et mes oncles Mourad et Mohamed,

A tous membre de ma famille,

A mon encadreur Mme Boulkrinat Nour El Houda, qui m'a donnée l'effort et le soutien moral pour la réalisation de ce projet.

A mon promotrice Mme Benblidia Nadjia

A Mon binôme durant cette épreuve ihcene, Pour ta sincérité. Je te souhaite une vie pleine de joie et une carrière de succès et réussite.

A tous mes amis de promotion de 2^{ème} année Master en IL,

A mes meilleures amies proches.

A tous ceux qui m'ont aidé de près ou de loin.

Dalila

Résumé

De nos jours, les réseaux sociaux apparaissent comme un nouveau moyen important de communication. Les réseaux sociaux sont un ensemble d'entités sociales où les utilisateurs sont reliés entre eux par des relations et interactions sociales (liens entre membres, informations qui circulent entre eux) pour partager des intérêts communs. Cependant, les intérêts des utilisateurs évoluant au fil du temps, la détection des changements des centres d'intérêts sociaux est liée par le comportement et l'interaction des utilisateurs dans leur environnement social.

Nous proposons dans ce travail une solution à l'évolution des centres d'intérêts des utilisateurs non actifs (profils passifs). À cet effet, nous avons dans un premier temps effectué la pondération des requêtes pour extraire les mots avec leur poids, ensuite la sélection des tops mots qui représenteront les centres d'intérêts de l'utilisateur. Par la suite, ces intérêts seront enrichis par l'ajout de nouveaux centres d'intérêts en calculant la similarité entre le profil non actif et les profils existants dans la base de données.

Mots-clés : Réseau social, profil social, utilisateur non actif, évolution du profil social, pondération temporelle, mesure de similarité.

Abstract

Nowadays, social networks are emerging as an important new means of communication. Social networks are a set of social entities where users are linked together through social relationships and interactions (links between members, information that circulates between them) to share common interests. However, as user interests evolve over time, the detection of changes in social interests is linked by the behavior and interaction of users in their social environment.

We propose in this work a solution to the non-active user's interest's evolution (passives profiles). To this end, we first carried out the requests weighting in order to extract the words and their weight, next the selection of the top words that will represent the user's interests. Afterward, these interests will be enriched by adding new interests through the computation of the similarity between the inactive profile and the existing profiles in the database.

Keywords: Social profile, social network, user not active, evolution of the social profile, time weighting, similarity measure.

المخلص

الشبكات الاجتماعية كوسيلة جديدة مهمة للاتصال. الشبكات الاجتماعية هي مجموعة من الكيانات الاجتماعية حيث يتم ربط المستخدمين معا من خلال العلاقات الاجتماعية و التفاعلات (الروابط بين الأعضاء و المعلومات التي يتم تداولها بينهم) لمشاركة الاهتمامات المشتركة. ومع ذلك، مع تطور إهتمامات المستخدمين بمرور الوقت، يرتبط إكتشاف التغيرات في الإهتمامات الاجتماعية بسلوك وتفاعل المستخدمين في بيئتهم الاجتماعية.

في هذا العمل ، نقترح حلا للمصالح المتغيرة للمستخدمين غير النشطين (ملفات التعريف السلبية) . ولهذه الغاية، قمنا أولا بترجيح الطلبات لإستخراج الكلمات مع وزنها ، تم اختيار أفضل الكلمات التي ستمثل مراكز إهتمام المستخدم .

بعد ذلك ، سيتم إثراء هذه الاهتمامات عن طريق إضافة اهتمامات جديدة عن طريق حساب التشابه بين الملف الشخصي غير النشط والملفات الشخصية الموجودة في قاعدة البيانات .

الكلمات المفتاحية : الشبكة الاجتماعية ، الملف الشخصي الاجتماعي ، المستخدم غير النشط ، تطور الملف الاجتماعي، وزن الوقت ، قياس التشابه.

Table des matières

Introduction générale.....	1
----------------------------	---

Chapitre 1 : Notions fondamentales sur les réseaux sociaux

1. Introduction	3
2. Réseau social	3
3. Types des réseaux sociaux	4
3.1. Classiques / Numériques	4
3.2. Comparaison.....	4
4. Réseau sociaux numériques	4
4.1. Types	4
4.2. Spécification des réseaux sociaux numériques	5
5. Statistiques sur les réseaux sociaux.....	6
6. Construction de réseau social	7
7. Présentation et éléments d'un réseau social	7
7.1. Nœuds.....	8
7.2. Liens	8
7.2.1. Orientation des liens	8
7.2.2. Pondération des liens	8
7.2.3. Réseau multidimensionnel.....	9
7.3. Groupes	10
7.4. Tags	10
7.5. Ressources (Contenus)	10
7.6. Graphe de contenu social	10
8. Conclusion	11

Chapitre 2 : Profilage et évolution des intérêts sociaux

1. Introduction	12
2. Profilage social	12
3. Profil social et données sociales.....	13
4. Types d'utilisateurs	13
5. Profil non actif.....	13

6. Modélisation du profil social.....	14
6.1. Construction du profil social.....	14
6.2. Représentation du profilage	15
6.3. Evolution du profil social.....	16
7. Les techniques d'évolution du profil social	17
7.1. Evolution basée sur les personnes proche / tags / ressources	17
7.1.1. Détection des intérêts basée sur les personnes proches.....	17
7.1.2. Détection des intérêts basée sur les tags.....	17
7.1.3. Détection des intérêts basée sur les ressources	18
7.2. Evolution temporelle.....	18
7.3. Algorithmes évolutifs.....	19
7.3.1. Algorithme génétique	19
7.3.2. Réseau de neurones	20
8. Conclusion.....	21

Chapitre 3 : Conception du système d'évolution des profils sociaux

1. Introduction	22
2. Architecture du système	22
2.1. Réseau social.....	23
2.2. Préparation des données.....	23
2.2.1. Indexation des ressources	23
2.2.2. Extraction des données	24
2.2.3. Prétraitement.....	25
2.2.4. Jeu de données	25
2.3. Espace utilisateur	26
2.4. Espace administrateur	26
3. Evolution du profil social	27
3.1. Pondération des requêtes.....	28
3.1.1. La pondération temporelle.....	28
3.1.2. Poids des mots clés	29
3.1.3. Top mots clés.....	32
3.2. Mesure de similarité.....	32
4. Diagrammes UML représentatifs du système	34
4.1. Diagramme des cas d'utilisation.....	35

4.2. Diagramme de séquence	36
4.2.1. Diagramme de séquence du cas « S’authentifier ».....	36
4.2.2. Diagramme de séquence du cas « Traiter l’évolution »	37
4.3. Diagramme de classes.....	38
4.4. Passage du diagramme d’objet vers le modèle relationnel	39
5. Conclusion.....	40

Chapitre 4 : Mise en œuvre et expérimentation du système d’évolution des profils sociaux

1. Introduction	41
2. Environnement de développement	41
2.1. Python.....	41
2.2. Spydar.....	42
2.3. PyQt.....	42
3. Interfaces de l’application Environnement de développement	42
3.1. Authentification.....	42
3.2. Espace utilisateur.....	43
3.2.1. Consultation des requêtes.....	44
3.2.2. Consultation statistiques.....	44
3.3. Espace administrateur	45
3.3.1. Visualisation des profils	46
3.3.2. Evolution des profils sociaux	47
3.3.3. Statistiques	48
4. Conclusion.....	49
Conclusion générale	50
Bibliographie.....	51

Liste des Figures

Chapitre 1 : Notions fondamentales sur les réseaux sociaux

Figure 1 : Quelques types des réseaux sociaux numériques	5
Figure 2 : Les réseaux sociaux les plus populaires en 2021.....	6
Figure 3 : Présentation un réseau social par un graphe et une matrice	7
Figure 4 : Représentation graphique réseau non orienté et non pondéré à trois nœuds (a) et réseau orienté et pondéré à trois nœuds (b).....	9
Figure 5 : Réseau multidimensionnel.....	9

Chapitre 2 : Profilage et évolution des intérêts sociaux

Figure 6 : Opérateur de croisement / mutation	19
Figure 7 : Architecture d'un réseau de neurones.....	20

Chapitre 3 : Conception du système d'évolution des profils sociaux

Figure 8 : Architecture du système	22
Figure 9 : Extraction des données à partir du réseau social Facebook.....	25
Figure 10 : Schéma conceptuel des informations résultantes	26
Figure 11 : Processus d'évolution du profil social	28
Figure 12 : Diagramme des cas d'utilisation.....	35
Figure 13 : Diagramme de séquence du cas « S'authentifier ».....	37
Figure 14 : Diagramme de séquence du cas « Traiter l'évolution »	38
Figure 15 : Diagramme de classes	39

Chapitre 4 : Mise en œuvre et expérimentation du système d'évolution des profils sociaux

Figure 16 : Les outils d'environnement de développement	41
Figure 17 : Page authentification	43
Figure 18 : Espace utilisateur	43
Figure 19 : Consultation l'historique des recherches.....	44

Figure 20 : Consultation statistiques des centres d'intérêts	45
Figure 21 : Espace administrateur	46
Figure 22 : Visualisation des profils	47
Figure 23 : Evolution des profils sociaux	48
Figure 24 : Statistiques des poids des centres d'intérêts	49

Liste des algorithmes et des tableaux

Chapitre 3 : Conception du système d'évolution des profils sociaux

Algorithme 1 : Calcul des poids des mots de chaque utilisateur	31
Algorithme 2 : Fonction de similarité	34
Tableau 1 : Description des rôles des acteurs	36

Introduction générale

Introduction générale

Contexte

Les réseaux sociaux ont eu un grand succès ces dernières années, avec des millions d'utilisateurs visitant des sites tels que Facebook, Twitter, Flickr, Delicious, ...etc. Ces sites de médias sociaux s'appuient principalement sur leurs utilisateurs pour la création du contenu, annotation du contenu des autres, et établissement des relations en ligne. Les activités des utilisateurs reflètent leurs opinions, intérêts, ... etc. dans leur environnement social. Suivre l'évolution des intérêts des utilisateurs permet de répondre rapidement aux besoins d'un utilisateur ou d'un groupe d'utilisateurs partageant les mêmes domaines d'intérêts. Cette évolution repose sur la détection implicite des centres d'intérêts à travers les comportements observables collectés lors de l'interaction de l'utilisateur avec son réseau social et les commentaires (tags) sur les publications (textes, vidéo, images, ...) partagées entre amis.

Les intérêts d'utilisateurs évoluant au fil du temps, il en est de même pour ceux extraits depuis son réseau social : pertinents à un moment donné, ils peuvent ne plus être significatifs ultérieurement. La détection des intérêts des utilisateurs actifs se base sur les interactions avec les réseaux sociaux (partager, commenter, aimer, ...etc.) sur les documents (texte, image, vidéo, ...). Cependant, détecter les intérêts d'un utilisateur non actif et suivre son évolution est problématique vu que ce type d'utilisateur n'interagit pas (ou rarement) avec les réseaux sociaux et ne possède aucun ami (ou très peu). Les facteurs de son inactivité peuvent être des causes de santé, familiales ou des causes dangereuses comme des faux profils, ou des comptes malveillants (piratage). A notre connaissance (après plusieurs recherches) il n'existe pas des travaux qui ont déjà traité l'évolution des profils non actifs.

Objectifs

Dans notre travail, nous proposons de résoudre le problème d'évolution des intérêts des utilisateurs non actifs, cette évolution se basera sur l'analyse de différentes informations recherchées par l'utilisateur dans son réseau social (requêtes) ainsi que son interaction avec le système (consultation des ressources). La réalisation du système d'évolution des profils sociaux se fera selon plusieurs étapes :

- Maitrise du contexte des réseaux sociaux.
- Etude des approches existantes pour l'évolution du profil social.
- Synthèse et choix d'une approche pour l'évolution du profil social des utilisateurs non actifs.

- Proposition d'une solution permettant de faire le profilage et l'évolution des utilisateurs non actifs.
- Développement d'une application desktop de suivi d'évolution des utilisateurs non actifs sous environnement python.

Organisation de mémoire

Nous structurons ce présent mémoire en quatre chapitres afin d'aborder les principales étapes du projet. Il est organisé comme suit :

Chapitre 1 : représente les concepts fondamentaux du réseau social, ses éléments et ses propriétés.

Chapitre 2 : consacré au profil social, types d'utilisateurs et à l'évolution des intérêts sociaux.

Chapitre 3 : décrit notre démarche architecturale et conceptuelle pour la mise en œuvre de notre solution.

Chapitre 4 : illustre la concrétisation et la réalisation de notre système et la représentation des résultats obtenus.

Enfin, ce mémoire se termine par une conclusion générale.

Chapitre 1

*Notions fondamentales sur les
réseaux sociaux*

1. Introduction

La notion de réseau social a été introduite dans les années 1950 en sciences sociales par l'anthropologue britannique Barnes 1954 [Girard, 2012]. Depuis quelques années, le terme réseaux sociaux évoque souvent les réseaux sociaux numériques tels que Facebook, Twitter, ou YouTube, etc. Les réseaux sociaux ont vu le jour aux Etats-Unis peu de temps après la propagation d'Internet dans les années 1990. L'explosion de la création de nouveaux réseaux sociaux a eu lieu grâce au développement des téléphones portables¹.

Les réseaux sociaux sont devenus un moyen de communication autonome et libre [Cross et al., 2004]. L'utilisation des réseaux sociaux devient de plus en plus usuelle, elle est devenue un phénomène en vogue de nos jours et le nombre d'utilisateurs ne fait qu'augmenter. Ils prennent une place de plus en plus importante dans la vie personnelle et professionnelle.

Les réseaux sociaux ne sont en effet qu'une partie des médias sociaux, les médias sociaux désignent l'ensemble des services permettant de développer des conversations et des interactions sociales sur internet ou en situation de mobilité [Cavazza, 2009].

2. Réseau Social

Plusieurs définitions ont été introduites dans la littérature pour définir les réseaux sociaux, nous citons :

Les réseaux sociaux sont des services Web qui permettent aux individus de construire un profil public ou semi-public dans le cadre d'un système délimité, d'articuler d'une liste d'autres utilisateurs avec lesquels ils partagent des relations ainsi que de voir et croiser leurs listes de relations et celles faites par d'autres à travers la plateforme [Boyd et al., 2007].

Un réseau social est constitué d'organisations ou d'individus reliés entre eux par les liens qui sont créés à l'occasion d'interactions sociales [Fayon, 2008].

Le réseau social se définit comme une plate-forme de communication en ligne permettant à ses utilisateurs de partager des intérêts communs et d'interagir entre eux. Il est en général proposé aux membres de mettre en ligne leur propre contenu ou ceux de tiers, tel que des photos, des vidéos, des textes, des liens ou des commentaires, qui seront accessibles et visibles par d'autres utilisateurs [Sargsyan, 2017].

¹ https://www.ac-orleans-tours.fr/fileadmin/user_upload/0281077U/Vie_du_lycee/Historique_des_reseaux_sociaux.pdf

3. Types des réseaux sociaux

3.1. Classiques / Numériques

Les réseaux sociaux peuvent être classés en deux catégories :

- ❖ **Réseau social traditionnel (classique)** : est défini comme une entité constituée d'un ensemble d'individus et des relations qu'ils entretiennent les uns avec les autres, directement ou indirectement par le biais de chaînes de relations, il peut aussi s'agir d'organisation, d'institution [Fondeur et al., 2006].
- ❖ **Réseau social numérique (RSN)** : est un terme général utilisé pour décrire un ensemble de technologies basé sur le Web. Il désigne une panoplie de site Web ou de plateformes numérique qui abondent sur le Web 2.0 [Mercier, 2008].

3.2. Comparaison

Les diverses différences entre les réseaux sociaux numériques (en ligne) et les réseaux sociaux traditionnels (sans internet) sont :

- ❖ Le réseau social numérique contient un grand nombre de liens par contre le réseau social traditionnel contient un petit nombre de liens.
- ❖ Le réseau social numérique possède des liens faibles et forts mais le réseau social traditionnel possède des liens forts seulement.
- ❖ Le réseau social traditionnel contient des personnes connues, par contre le réseau social numérique contient des personnes connus et inconnus.
- ❖ Les personnes influentes sur les deux réseaux sociaux numériques et traditionnels.
- ❖ Le réseau social numérique se base sur le Web (internet) par contre le réseau social traditionnel en dehors d'internet.

4. Réseaux sociaux numériques

4.1. Types

Les réseaux sociaux numériques sont classés en différents types (**Figure 1**) :

- ❖ **Généraliste** : est un réseau social numérique de contact qui permette de créer, d'agrandir son cercle d'amis et discuter, comme : Facebook², Twitter³.

² www.facebook.com

³ www.twitter.com

- ❖ **De partage** : est un réseau social numérique de contenu qui permette de mettre en ligne, partager des vidéos, photos, liens et musiques par exemple : YouTube⁴, Flickr⁵, Delicious⁶.
- ❖ **Professionnel** : est un réseau social numérique de contact qui permette de construire des réseaux sociaux professionnels personnalisés (trouver un nouvel emploi, annoncer des évènements ou des activités professionnelles), par exemple : LinkedIn⁷



Figure 1 : Quelques types des réseaux sociaux numériques

4.2. Spécifications des réseaux sociaux numériques

Selon [Joseph, 2014], les spécifications des réseaux sociaux sont :

- ❖ Multiplicité des plateformes.
- ❖ Gratuité et ouverture pour la plupart des plateformes.
- ❖ Principe d'invitation.
- ❖ Partage et possibilité d'exploitation de données personnelles.
- ❖ Interaction entre les membres.
- ❖ Partage d'applications.
- ❖ Regroupement des personnes autour de centres d'intérêts.
- ❖ Partage d'opinions et de sentiments.
- ❖ Structuration identitaire des membres.
- ❖ Possibilité de recrutement.
- ❖ Publicité.

⁴ www.youtube.com

⁵ www.flickr.com

⁶ www.delicious.com

⁷ www.linkedin.com

5. Statistiques sur les réseaux sociaux

En raison de la présence constante dans la vie de leurs utilisateurs, les réseaux sociaux ont un très grand impact social. Le flou entre le réel et le virtuel, ainsi que le concept d'interactions sociales en ligne sont des aspects qui ont émergé au cours de ces dernières années et sont devenus un outil indissociable de la vie des différentes tranches sociales : les jeunes, les entreprises, les chercheurs, etc.

Selon statista⁸, les réseaux sociaux les plus populaires (**Figure 2**) affichent généralement un nombre très élevé de comptes d'utilisateurs actifs.

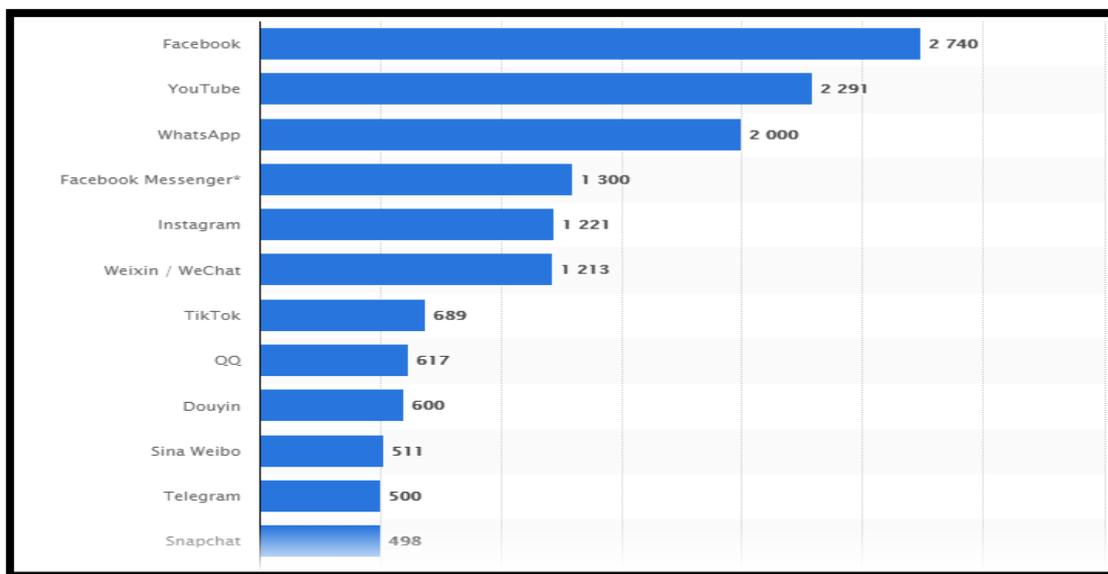


Figure 2 : Les réseaux sociaux les plus populaires en 2021 [Statista, 2021]

La **figure 2** montre des résultats mensuels des réseaux sociaux les plus populaires dans le monde en 2021, classés par nombre de comptes actifs. Cette dernière indique que Facebook domine avec un nombre total d'utilisateurs en hausse pour atteindre près de 2.8 millions utilisateurs. D'autres réseaux sociaux tels que YouTube et WhatsApp ont enregistré respectivement un taux de 2.2 millions et de 2.0 millions d'utilisateurs actifs.

6. Construction de réseau social

Deux catégories de construction de médias sociaux ont été proposées par [Kaplan et al., 2010] : implicite et explicite.

⁸ <https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/>

- ❖ Réseau social est dit **implicite** lorsque les liens entre utilisateurs peuvent être construits à partir des interactions ou actions des utilisateurs.
- ❖ Réseau social est dit **explicite** lorsque les liens entre utilisateurs sont construits explicitement par eux.

7. Présentation et éléments d'un réseau social

L'analyse des réseaux sociaux permettent de formaliser les interactions sociales en termes des nœuds et des liens. Le réseau social est présenté par un graphe $G = (V, E)$ où V est l'ensemble des nœuds représentant les entités sociales (acteurs sociaux) et E est l'ensemble des associations entre les nœuds dans V .

Un graphe peut être représenté à l'aide une matrice dite adjacence (A_{ij}) qui indique la connexion entre les nœuds. Traditionnellement, la valeur numérique **0** marquant l'absence de connexion, **1** marquant la présence de connexion.

La **figure 3** illustre la présentation d'un réseau social.

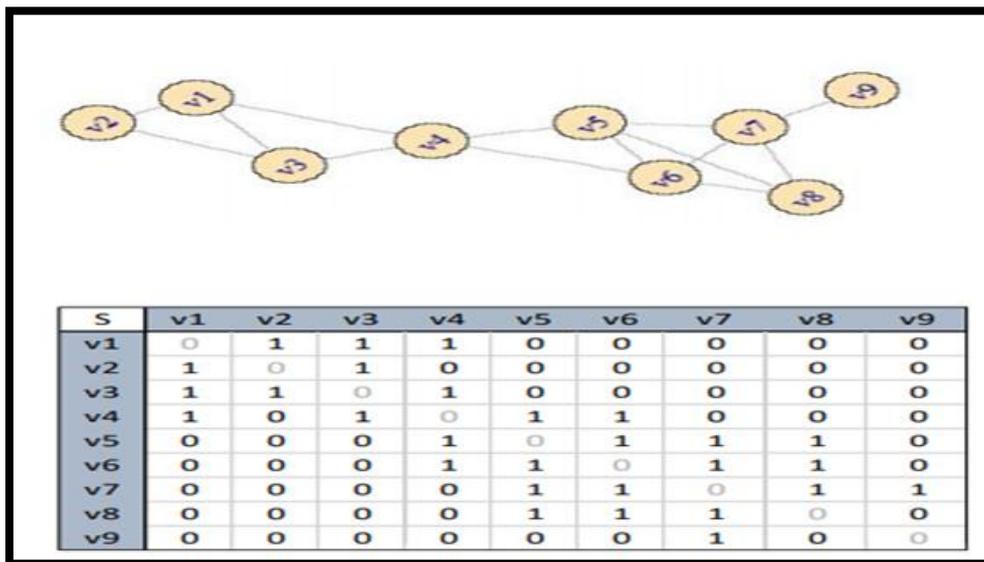


Figure 3 : Présentation d'un réseau social par un graphe et une matrice [Tang et al., 2010]

Les éléments d'un réseau social sont :

7.1. Nœuds

Dans la théorie des graphes les acteurs sont représentés par des nœuds ou sommets. Chaque nœud possède des attributs individuels de types différents : simples (numériques par

exemple : âge, textuels par exemple : la taille, sexe) ou structurés comme des vecteurs (par exemple : intérêts).

7.2. Liens

Les liens sont des relations entre ces acteurs (nœuds) appelées aussi liens ou arrêtes. Elles permettent d'associer des nœuds par paire.

Les relations sociales qui caractérisent un réseau peuvent être définies par différents traits comportementaux, catégorisés sous les termes relations et interactions [Sueur et al., 2011].

Le terme « relation » décrit deux ou plusieurs individus qui sont en même endroit au même moment (association, contact et proximités entre individus). Par contre le terme « interaction » est défini comme un comportement dirigé d'un individu émetteur vers le récepteur par exemple envoyé des messages. Les propriétés des liens sont :

7.2.1. Orientation des liens

Les liens dans les réseaux sociaux sont classés en deux types : réseau orienté (dirigé) et réseau non orienté (non dirigé). La **figure 4** ci-dessus illustre les types de réseau.

- ❖ **Réseau orienté** : les liens dans un réseau dirigé sont orientés d'un émetteur vers récepteur sont asymétriques. Le réseau dirigé représente des interactions entre les individus.
- ❖ **Réseau non orienté** : les liens dans un réseau non dirigé sont symétriques. Les associations spatiotemporelles (contact, proximité entre individus) sont symétriques et le réseau qui les représente est non dirigé.

7.2.2. Pondération des liens

Un réseau peut également être qualifiés de pondérer et non pondéré (**Figure 4**), tel que :

- ❖ **Réseau pondéré (valorisé)** : représente des relations de force différente quand la fréquence ou la durée des interactions peut être enregistrés.
- ❖ **Réseau non pondéré** : qui possède une valeur binaire est attribué à chaque lien. Si aucune interaction n'a été observée (pas de lien), il a une valeur 0, sinon 1 lorsque au moins une interaction a été observé entre deux individus.

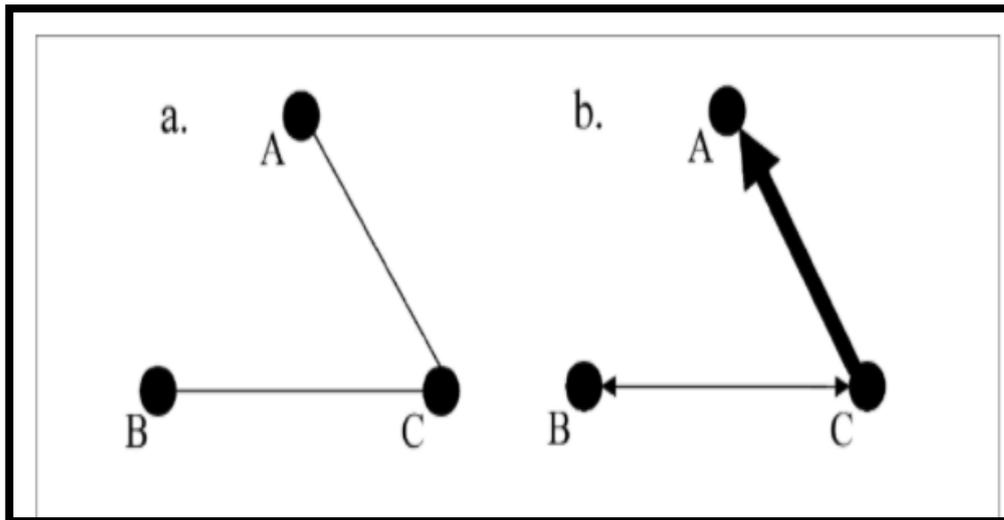


Figure 4 : Représentation graphique réseau non orienté et non pondéré à trois nœuds (a) et réseau orienté et pondéré à trois nœuds (b)

7.2.3. Réseau multidimensionnel

C'est un réseau qui permet de définir plusieurs types de relations entre deux ou plusieurs individus [Berlingiero et al., 2013]. Un réseau multidimensionnel (**Figure 5**) est un type particulier de réseau multicouche (plusieurs réseaux évoluant et interagissant entre eux), il est représenté par un triple où se trouve un ensemble de dimensions (couches), dont chaque membre est un type de lien différent et se compose de triplet, tel que : $G = (V, E, D)$.

V : ensemble des nœuds (individus).

E : ensemble des liens (connexions).

D : ensemble des dimensions.

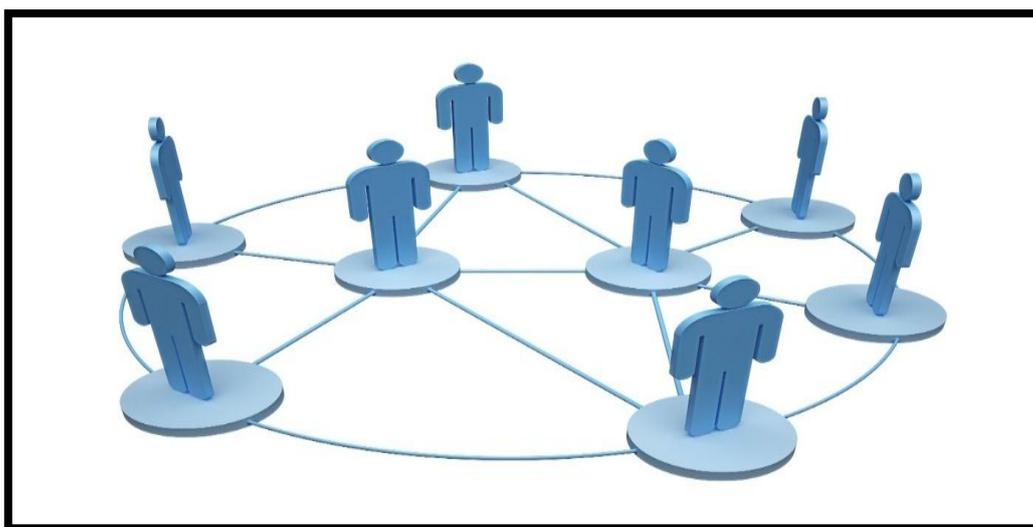


Figure 5 : Réseau multidimensionnel

7.3. Groupes

Un Groupe (ou cluster) est défini comme un ensemble de nœuds fortement connectés entre eux et plus faiblement connectés au reste du réseau.

Ces groupes de nœuds (composantes) sont appelées « communautés d'intérêt ». Une communauté peut être définie, d'une manière structurelle comme un ensemble de nœuds qui ont plus de liens entre eux qu'avec les autres groupes, et d'une manière sémantique comme un ensemble de nœuds partageant les mêmes centres d'intérêt.

7.4. Tags

Un tag (annotation) est un mot clé ou terme associé à une information (image, clip vidéo ou article) qui décrit une caractéristique de l'objet et permet un regroupement facile des informations contenant les mêmes mots clés⁹.

Un tag peut être :

- ❖ **Tag personnel** : reflète le sentiment de l'utilisateur et non pas le contenu de la ressource.
- ❖ **Mot propre** : qui n'est pas compréhensible par les autres utilisateurs et le système.
- ❖ **Spam** : promouvoir un intérêt d'un autre utilisateur.

7.5. Ressources (Contenus)

Les ressources sur les médias sociaux peuvent être de n'importe quel type : URL, vidéo, image, etc., que l'utilisateur peut commenter (tag), partager, aimer...etc.

7.6. Graphe de contenu social

D'après [Amer-Yahia et al., 2007], un graphe de contenu social est composé de deux types de nœuds sont les individus (utilisateurs) et les contenus (comme : texte, vidéo, photo). Les relations entre les nœuds varient en fonction de type du nœud source et nœud cible, nous distinguons quatre types de relation :

- ❖ **Personne-Personne** : présente les relations sociales entre les personnes, par exemple : amis, familles, collègue de travail.
- ❖ **Personne-Contenu** : désigne la relation entre les individus et les contenus, par exemple partager, commenter, publier et aimer.
- ❖ **Contenu-Contenu** : présente la relation entre les contenus, par exemple : deux commentaires sur un contenu, plusieurs aiment sur un contenu.

⁹ [https://fr.wikipedia.org/wiki/Tag_\(m%C3%A9tadonn%C3%A9e\)](https://fr.wikipedia.org/wiki/Tag_(m%C3%A9tadonn%C3%A9e))

- ❖ **Contenu-Personne** : désigne la relation d'un contenu vers individu, par exemple : un commentaire est écrit par individu.

8. Conclusion

A travers ce premier chapitre, nous avons détaillé le réseau social, tout d'abord nous avons défini le réseau social ainsi leurs types et la comparaison entre eux. Ensuite, nous avons expliqué le réseau social numérique, ses types, ses spécifications et sa construction. Par la suite, nous avons présenté les éléments et les propriétés d'un réseau social. Nous avons constaté que le nombre d'utilisateur augmente rapidement dans les réseaux sociaux. A cet effet, nous étudions dans le deuxième chapitre le profilage et l'évolution des intérêts sociaux des utilisateurs.

Chapitre 2

*Profilage et évolution des
intérêts sociaux*

1. Introduction

Les réseaux sociaux sont de plus en plus utilisés, les utilisateurs y échangent des informations et fournissent des éléments sur leur profil. Ces données peuvent être utilisées pour modéliser un individu selon les activités qu'il réalise sur le réseau social : il s'agit du profil de l'utilisateur [Ramiandrisoa et al., 2017].

[Zemirli, 2008] définit un profil d'utilisateur comme une structure qui permet de modéliser et de stocker les données caractérisant l'utilisateur. Ces données représentent les centres d'intérêts, les préférences et les besoins en informations de l'utilisateur ou un groupe d'utilisateurs [Zemirli et al., 2005], qui contient ses données personnelles, son historique, ses annotations et ses intérêts.

L'évolution du profil utilisateur social se fait en utilisant des techniques pour détecter les changements sur ces centres d'intérêts à travers le temps, la mise à jour est une partie de l'évolution qui consiste à modifier, ajouter ou supprimer des intérêts du profil. Elle permet d'extraire des nouvelles informations ou compléter les intérêts du profil.

2. Profilage social

Le profilage apparaît comme une technique particulière de traitement des données et consiste en une analyse des comportements, des actions et interactions d'un individu provenant de sources différentes [Walter, 2011]. Le but du processus du profilage est de déduire certaines caractéristiques ou révéler des informations non apparentes pour l'internaute mais qui lui sont inférées afin de lui attribuer un certain profil. [King et al., 2016] définit le profil comme « un ensemble de données qui caractérise un individu ou une catégorie d'individus et qui est destiné à être appliqué à un individu ». Ceci dans la perspective de prendre des décisions individuelles à son égard, de prédire ou d'influencer son comportement et ses choix.

Nous distinguons deux types de profilages :

- ❖ **Profilage abstrait** : n'est pas fondé sur des données à caractère personnel mais sur des données anonymes ou antonymiques et qui par conséquent, ne se rapporte pas à une personne identifiée ou identifiable [Meier, 2011].
- ❖ **Profilage spécifique** : résulte de l'analyse de données fournies par la personne concernée, sans inférence.

3. Profil social et données sociales

Les données sociales d'une personne fait référence aux données personnelles générées en ligne ou hors ligne. Une grande quantité de ces données, y compris la langue, la localisation et les intérêts est partagée via les médias sociaux et les réseaux sociaux. Globalement, ces informations peuvent construire le profil social d'une personne.

Le profilage social est une approche émergente visant à surmonter les difficultés rencontrées pour répondre aux demandes des utilisateurs en introduisant le concept de recherche personnalisée tout en tenant compte des profils d'utilisateurs générés à l'aide de données de leur réseau social. Les données des médias sociaux de Twitter et Facebook ont été utilisées par la plupart des études pour déduire les attributs sociaux des utilisateurs. La littérature a montré que les attributs sociaux des utilisateurs, notamment l'âge, le sexe, le lieu de résidence, le bien-être, les émotions, l'opinion, les relations, l'influence restaient à explorer [Bilal et al., 2019].

4. Types d'utilisateurs

À partir du profilage des utilisateurs, nous en distinguons trois types d'utilisateurs [Kumar et al., 2008] :

- ❖ **Membres passifs** : qui ne participent pas à l'évolution du réseau social.
- ❖ **Membres linkers** : sont les utilisateurs les plus actifs et sociables qui sont situés dans un noyau fortement connecté.
- ❖ **Membres inviters** : sont les utilisateurs qui encouragent leurs amis hors ligne et leurs à la participation dans les réseaux sociaux.

Les relations des utilisateurs avec les autres sont sur la plupart des plateformes, classées uniquement comme "amis" ou "non amis" et les utilisateurs peuvent transmettre des informations importantes aux "amis" de leur cercle social, mais pas nécessairement aux utilisateurs avec lesquels ils souhaitent consciemment partager les informations [Fogues et al., 2015].

5. Profil non actif

Selon les types d'utilisateurs, le profil non actif (passif) est considéré comme un profil qui interagit moins avec son réseau social, non spécifique et difficile à comprendre ses identités, le profil non actif peut être :

- ❖ **Compte malveillant** : il existe de nombreuses activités malveillantes sur les réseaux sociaux, notamment l'utilisation d'ingénierie d'attaque sociale et de logiciels malveillant de distribution de spams. Les spammeurs utilisent une stratégie d'ingénierie d'attaque sociale pour voler les informations d'identification d'utilisateurs légitimes et éventuellement compromettre leurs comptes [Egele et al., 2015]. Les comptes utilisés pour des activités malveillantes sont classés en deux catégories :
 - **Fraudes** : sont des comptes pour diffuser des contenus malveillants comme l'intégration des liens malveillants vers des pages Web afin d'obtenir des informations sensibles de la victime [Chen et al., 2014].
 - **Comptes compromis (piratés)** : sont des comptes détournés d'utilisateurs légitimes à l'aide une stratégie telle que l'ingénierie d'attaque sociale pour tromper les utilisateurs légitimes en cliquant sur des liens vers des pages Web. Les comptes compromis sont plus utiles pour les spammeurs que les comptes de spam de carrière, car ils permettent aux spammeurs de tirer parti de la relation de confiance existante entre les comptes et les fournisseurs de réseaux sociaux [Egele et al., 2015].
- ❖ **Faux profil** : contient des informations volées à des utilisateurs légitimes peuvent être utilisées pour créer des comptes Sybil¹⁰ afin de tromper les amis des vrais utilisateurs [Bilge et al., 2009], ou pour envoyer des messages de spam personnalisés [Fire et al., 2014 ; Ezpeleta et al., 2015].

6. Modélisation du profil social

La modélisation du profil social de l'utilisateur est au cœur de la mise en œuvre de processus d'accès personnalisé à l'information. Elle consiste à décrire les caractéristiques informationnelles des utilisateurs à travers un modèle de profil, plusieurs techniques furent développées dans la littérature pour modéliser l'utilisateur, elles diffèrent cependant dans l'approche de représentation et construction du profil, nous présentons les principales phases du processus de modélisation :

6.1. Construction du profil social

¹⁰ Compte Sybil est une attaque au sein d'un système de réputation qui est renversé par la création des fausses identités dans un réseau.

La construction d'un profil utilisateur social nécessite d'effectuer des analyses sur les données en fonction de la modélisation utilisateur mise en œuvre, la construction du profil peut être explicite et implicite.

- ❖ **Construction explicite** : est une technique simple et directe pour obtenir les informations de l'utilisateur, elle consiste à interroger l'utilisateur pour lui demander des informations personnelles, démographiques et ses intérêts [Gauch et al., 2007].
- ❖ **Construction implicite** : consiste à collecter les données de l'utilisateur, en observant son comportement et en scrutant son activité, les techniques de construction implicite du profil social de l'utilisateur sont classées en deux catégories :
 - **Approche basée sur les individus** : cette approche est basée sur les individus voisins, les intérêts d'un utilisateur donné sont extraits à partir des informations de ses voisins sociaux considérés de manière individuelle. À partir des informations de chaque individu du réseau social de l'utilisateur, les intérêts sont extraits et calculés en se basant sur les caractéristiques propres de chaque individu [Cabanac, 2011].
 - **Approche basée sur les communautés (groupes)** : est une approche basée sur les communautés, les intérêts d'un utilisateur sont extraits à partir de chaque communauté extraite depuis son réseau social. Les intérêts sont donc extraits et calculés en se basant sur les caractéristiques de chaque communauté [Tchuente et al., 2013].

6.2. Représentation du profilage

La construction d'un profil utilisateur nécessite un modèle de représentation du résultat produit, nous distinguons trois modèles de représentation : représentation ensembliste, sémantique et conceptuelle [Gauch et al., 2007].

- ❖ **Représentation ensembliste** : consiste à représenter le profil utilisateur par des vecteurs de termes ou mots clés pondérés. La pondération permet de moduler l'importance de chaque intérêt par rapport à tous les autres intérêts de l'ensemble. Chaque terme peut simplement représenter un intérêt.
- ❖ **Représentation sémantique** : cette représentation consiste à enregistrer les intérêts de l'utilisateur dans les nœuds représentant les termes intérêts et les liens entre nœuds représentant la proximité sémantique.

- ❖ **Représentation conceptuelle** : est similaire à la représentation sémantique. Elle contient des nœuds de termes reliés par des relations, cette représentation utilise la hiérarchie des concepts (termes ou mots clés).

6.3. Evolution du profil social

L'évolution des intérêts dans le profil social consiste à des changements de ses intérêts à travers le temps [Crabtree et al., 1998], le profil utilisateur est utilisé par des mécanismes tels que :

- ❖ **Le système de recommandation** : consiste à accomplir un filtrage d'information (une technique pour d'extraire des informations pertinentes et de qualités à partir d'une grande masse d'informations) afin de suggérer aux autres utilisateurs avec qui interagir (se connecter) [Ricci et al., 2011].
- ❖ **Le système de recherche d'information personnalisée** : qui permet de modéliser et stocker les informations de l'utilisateur.

La prise en compte de cette évolution peut s'effectuer lors de la modélisation du profil, les intérêts de l'utilisateur correspondent à une des caractéristiques les plus fréquentes dans un profil utilisateur surtout dans les systèmes de recommandation, elles peuvent être classés en deux types : les intérêts à court terme et à long terme [Benammar et al., 2002].

- ❖ **Profil à court terme** : se définit lorsque l'utilisateur est intéressé par le sujet durant une période limitée, l'intérêt soit éphémère ou change plus fréquemment.
- ❖ **Profil à long terme** : se définit lorsque l'utilisateur est toujours intéressé par le sujet et que cet intérêt ne change que très rarement.

Le nombre des centres d'intérêts de l'utilisateur augmente avec la connaissance de nouvelles données d'interactions, tandis que le plus souvent l'ensemble des centres d'intérêts reste intact. Ainsi, l'évolution consiste à adapter la structure et le contenu du profil aux changements des centres d'intérêts et aux variations des besoins en information de l'utilisateur.

Dans le cas d'une représentation ensembliste, le profil utilisateur évolue en ajoutant de nouveaux vecteurs de termes extraits des documents correspondant aux centres d'intérêts détectés de l'utilisateur, mais dans le cas d'une représentation en classes (hiérarchiques ou non), le processus de mise à jour consiste à mettre à jour graduellement le classifieur.

L'évolution du profil social de l'utilisateur se base sur des techniques de détection des intérêts à partir des utilisateurs proches, des tags et des ressources [Astrain et al., 2010], l'utilisation d'un critère temporelle pour cerner l'évolution et la dynamique des informations étudiées [Maloof et al., 2000 ; Cheng et al., 2008]. Aussi, des algorithmes d'évolution génétique [Moukas, 1997], réseau de neurones [Kuflik et al. 2006] et des principes d'évolution liées à la théorie de la vie artificielle [Chen et al., 2002]. Au meilleur de notre connaissance (après plusieurs recherches) il n'existe pas des travaux qui ont déjà traité l'évolution des profils non actifs. A cet effet, nous présentons dans la section suivante les travaux sur l'évolution des utilisateurs actifs.

7. Les techniques d'évolution du profil social

7.1. Evolution basé sur les personnes proches / tags / ressources

7.1.1. Détection des intérêts basée sur les personnes proches

Les personnes proches (utilisateurs similaires) sont considérées comme une source importante d'information pour la détection des intérêts. En effet, les informations issues des personnes proches ont prouvé leur utilité pour surmonter le problème du démarrage à froid pour les nouveaux utilisateurs du système, pour détecter les intérêts des utilisateurs [Tchunte et al., 2013 ; Canut et al., 2015] et aussi pour enrichir les profils utilisateurs dans un but de recommandation [Meo et al., 2010 ; Kim et al., 2011]. La définition d'une personne proche est la relation sociale de l'utilisateur avec d'autres utilisateurs. Cette relation peut être explicite (une relation d'amitié) ou implicite (les utilisateurs qui agissent sur le même document par exemple : publication, image, vidéo...etc.). Ces personnes proches sont détectées par plusieurs techniques, tel que : par des observations que deux personnes sont proches si elles partagent des tags communs et donc elles peuvent bien avoir des intérêts en commun [Kim et al., 2011]. Aussi par la proximité de la fréquence d'interaction entre les utilisateurs [Roth et al., 2010]. La proximité est calculée selon [Guy et al., 2010] par des critères : le nombre des personnes et/ou de tags dans le profil utilisateur qui sont liés à un article, le degré de connectivité de ces personnes et/ou des tags à un utilisateur, le degré de connectivité de ces personnes et/ou tags à un article et la fraîcheur d'un article.

7.1.2. Détection des intérêts basée sur les tags

Plusieurs recherches portent sur la détection des intérêts de l'utilisateur à partir de l'information sociale produite par les utilisateurs et particulièrement des tags. Ces derniers sont considérés comme une information puissante pour refléter l'opinion de l'utilisateur vis-à-vis d'une ressource [Meo et al., 2010] et aussi pour détecter les intérêts de l'utilisateur [Kim

et al., 2011]. Les travaux qui détectent des intérêts en analysant les annotations sociales (tags), analysent : les tags les plus récents [Zheng et al., 2011], les tags plus populaires [Godoy et al., 2008], l'historique des annotations [Wang et al., 2010], ou la sémantique des tags [Kim et al., 2011], ces analyses peuvent fournir des tags pertinents pour l'utilisateur.

7.1.3. Détection des intérêts basée sur les ressources

Les intérêts peuvent être déduits sur la base des ressources que l'utilisateur accède [Ma et al., 2011 ; White et al., 2009]. La ressource peut être de n'importe quel type (URL, vidéo, image, etc.). Selon [Ma et al., 2011] les intérêts des utilisateurs sont découverts par extraction et par analyse des mots-clés de chaque source (les sources peuvent être Facebook, LinkedIn, etc.). Les intérêts des utilisateurs sont découverts à partir de l'analyse du comportement de l'utilisateur c'est-à-dire l'historique de visite des ressources, le temps passé sur une page Web, etc... [White et al., 2009]. Plusieurs techniques pour détecter les intérêts comme : l'indexation des ressources les plus pertinentes en utilisant des fonctions $(TF*IDF)^{11}$, tel que : **TF** est la fréquence d'un terme dans un document (le nombre d'occurrence d'un terme dans un document) et **IDF** est la fréquence inverse du document (l'importance d'un terme dans toute la collection) selon une requête spécifique [Vallet et al., 2010], cette requête peut être un tag et le contenu des ressources annotées. Aussi, l'approche de recommandation dans le système à base de tags, elle combine le contenu et l'analyse des relations dans un modèle unique [Zhang et al., 2010].

7.2. Evolution temporelle

Les centres d'intérêts d'une personne sont amenés à évoluer dans le temps, [Li et al., 2013 ; Kacem et al., 2014] proposent d'appliquer une fonction temporelle pour pondérer les intérêts de l'utilisateur selon leur fraîcheur, la détection des changements de centres d'intérêts sur les principales interactions n'est pas ciblée sur l'utilisateur lui-même mais sur les éléments de son réseau social (liens entre les membres, informations qui circulent entre les membres), les informations temporelles sont importantes pour l'évolution du profil social de l'utilisateur. Le critère temporel se base en deux facteurs : la dynamique des interactions entre les individus (utilisateurs) dans le réseau social et la dynamique des informations partagés [Spiliopoulou, 2011]. Les critères temporels sont considérés dans la phase de pondération des intérêts de l'utilisateur, le poids temporel est intégré dans la pondération du travail [On-At, 2017]. Cette proposition contient deux parties : la **première partie** consiste à attribuer aux informations, un poids prenant en compte le critère temporel (poids temporel) qui correspond à la prise en

¹¹ $TF*IDF$ est le produit des deux mesures TF et IDF.

compte de la dynamique des interactions entre les utilisateurs et des informations partagés entre eux, et pour la **deuxième partie** consiste à exploiter ce poids temporel, lors de l'étape de construction du profil sociale de l'utilisateur. Il s'agit de calculer la pertinence d'un individu par rapport à l'utilisateur central (un utilisateur centralise dans un ensemble des individus) en prenant en compte les informations temporelles de ses liens avec l'utilisateur central (date de création de liens avec l'utilisateur central, durée de la relation, ...etc.).

7.3. Algorithmes évolutifs

Les algorithmes évolutifs sont des outils utilisés dans divers domaines. Plusieurs algorithmes existent, nous distinguons deux algorithmes exploités dans le domaine de la recherche d'information : algorithme génétique et réseau de neurones.

7.3.1. Algorithme génétique

Algorithme génétique (AG) est développé par [Holland, 1992]. Un algorithme génétique est une algorithme stochastique itératif qui opère sur des ensembles de points codés, à partir d'une population initiale, et qui est construite à l'aide de trois opérateurs : croisement, mutation, et sélection qui sont la base de l'algorithme génétique. Le principe de fonctionnement s'apparente à celui du génome humain : les données sont converties en chaînes binaires, celles-ci se combinent par sélection, croisement, ou mutation, et donnent ainsi une nouvelle chaîne qui est évaluée. Ainsi ce processus est répété jusqu'à l'obtention de la solution pour le problème [Renders, 1995]. La **figure 6** ci-dessus représente les opérateurs de croisement ou mutation.

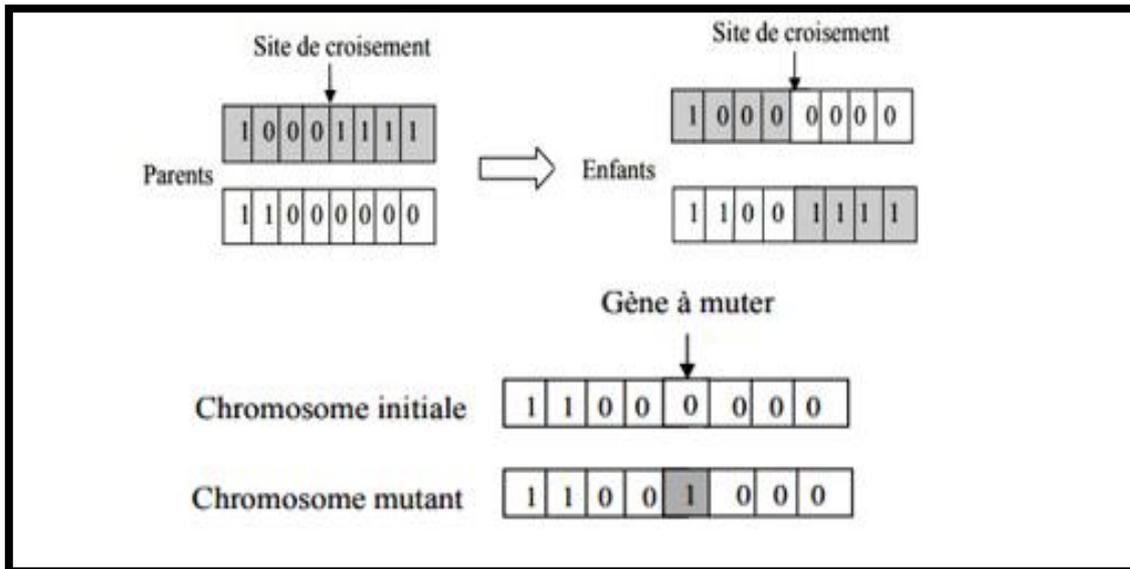


Figure 6 : Opérateur de croisement / mutation [Renders, 1995]

L'Algorithme génétique est utilisé dans le domaine de la recherche d'information, il a été exploité dans le travail de [Boulkrinat et al., 2019] pour l'enrichissement des centres d'intérêts en fonction des requêtes. Les mots des requêtes et des anciens intérêts avec leurs poids représentent l'ensemble de gènes qui composera par la suite les chromosomes. Ces derniers subiront des opérations génétiques afin de sélectionner le meilleur chromosome qui représentera les nouveaux centres d'intérêts de l'utilisateur.

7.3.2. Réseau de neurones

Les réseaux de neurones sont des outils très utilisés pour la classification, l'estimation et la prédiction. Ils sont issus de modèles biologiques, sont constitués d'unités élémentaires (les neurones) organisées selon une architecture comme l'illustre la figure 7 [Tufféry, 2012].

Un réseau de neurones est un ensemble de méthodes d'analyse et de traitements des données permettant de construire un modèle de comportement à partir de données qui sont des exemples de ce comportement. Un réseau de neurones est constitué d'un graphe pondéré orienté dont les nœuds symbolisent les neurones. Ces neurones possèdent une fonction d'activation qui permet d'influencer les autres neurones du réseau. Les connexions entre les neurones, que l'on nomme liens synaptiques, propagent l'activité des neurones avec une pondération caractéristique de la connexion [Houmadi, 2007].

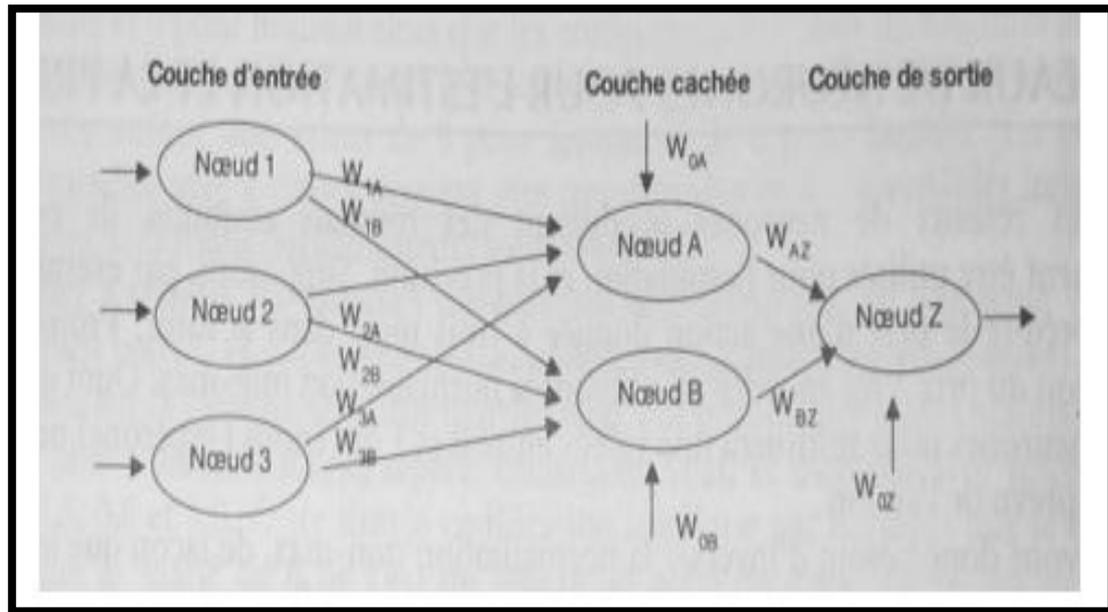


Figure 7 : Architecture d'un réseau de neurones [Tufféry, 2012]

L'intérêt du réseau de neurones a été présenté par [Kenter et al., 2018], où ils ont cité plusieurs domaines de son application : système de recommandation, modélisation du comportement des utilisateurs, ...etc. Il est aussi utilisé pour extraire l'information à partir d'un texte en langage naturel [Gupta, 2019].

8. Conclusion

A travers ce deuxième chapitre, nous avons présenté le profil social, types utilisateurs plus précisément le profil non actif, ensuite nous avons détaillé l'évolution du profil social et les techniques existantes pour détecter les intérêts de l'utilisateur dans le réseau social.

Dans le troisième chapitre, nous présentons la conception de notre système.

Chapitre 3

*Conception du système
d'évolution des profils sociaux*

1. Introduction

Nous concluons à travers les chapitres précédents que les intérêts des utilisateurs ont montré leur importance croissante dans la conduite du développement d'applications personnelles centrées sur l'utilisateur. Les études existantes pour analyser les intérêts d'utilisateurs se concentrent sur l'analyse de son interaction dans le réseau social et le contenu de sa navigation. L'un des problèmes les plus difficiles à résoudre est la détection des intérêts des utilisateurs peu actifs qui se connectent souvent mais interagissent pas sur leurs réseaux sociaux. A cet effet, nous proposons dans notre travail une solution qui permet l'évolution du profil social passif (peu actif) par la détection de ses centres d'intérêts à travers l'historique de ses recherches et les ressources visitées (consultées).

2. Architecture du système

La **figure 8** représente l'architecture globale de notre système qui s'articule autour de la notion du profil social et les différents éléments qui le constituent.

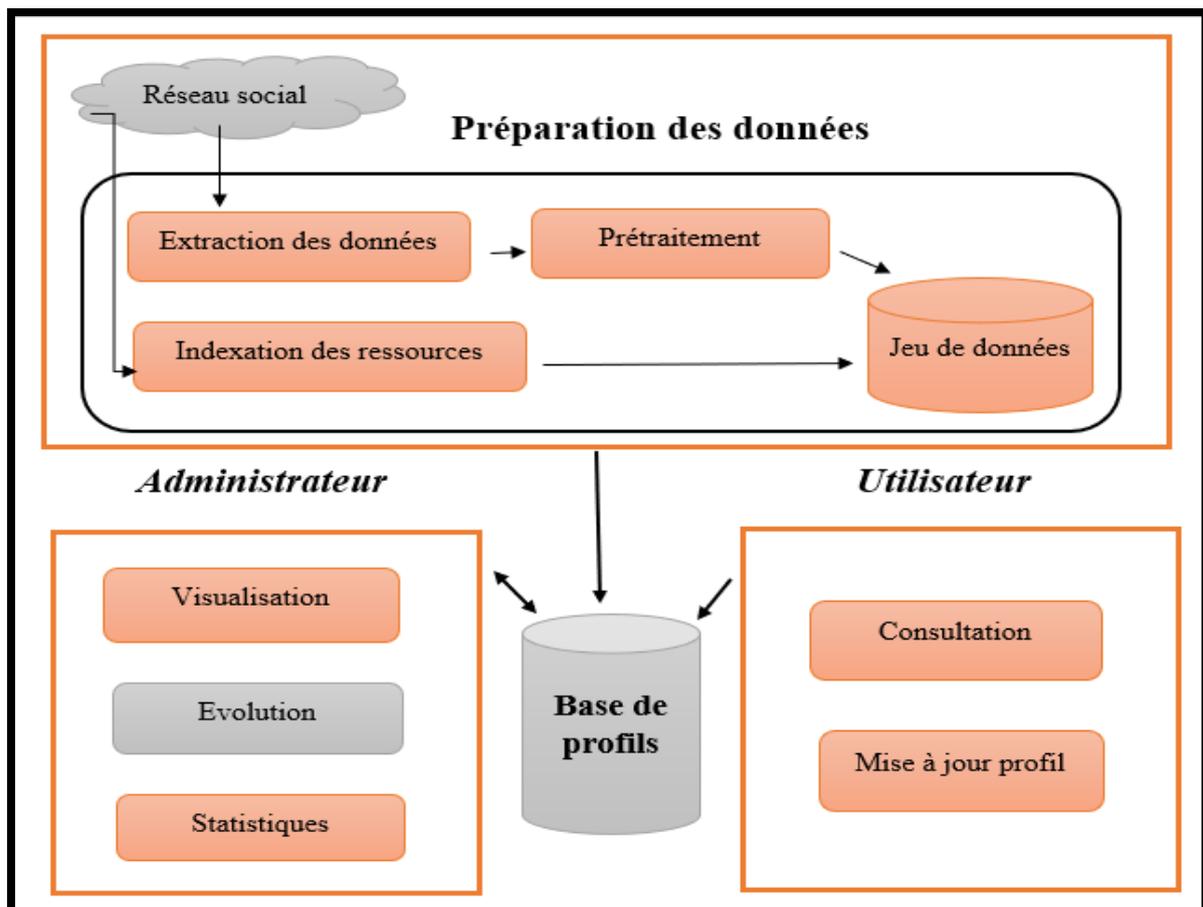


Figure 8 : Architecture du système

Nous identifions dans cette architecture, quatre principales composantes : réseau social, préparation des données, espace utilisateur, et espace administrateur.

2.1. Réseau social

Nous avons effectué notre étude à partir des données extraites du réseau social Facebook¹². Le réseau social Facebook¹³ est le plus gros réseau social du moment avec plusieurs centaines de millions de personnes inscrites. Il permet à ses utilisateurs de publier des images, des vidéos, des fichiers et documents, d'échanger des messages, joindre et créer des groupes et d'utiliser une variété d'applications sur une variété d'appareils. Vu l'inactivité de l'utilisateur, nous nous intéressons dans notre travail aux éléments social suivants :

- ❖ **Ressources** : publications textuelles, images, vidéos, consultés par l'utilisateur.
- ❖ **Recherche** : ensemble de requêtes (mots clés) exprimant le besoin d'information social (publications, pages, groupes publics).

2.2. Préparation des données

2.2.1. Indexation des ressources

L'indexation permet d'extraire tous les mots clés d'un document (dans notre cas ressource) et les présenter sous forme d'index. Les mots clés représentent le contenu de chaque ressource, où chaque mot clé possède un poids.

Pour le calcul du poids, nous allons utiliser la mesure **TF-IDF** [Vallet et al., 2010], comme suit :

La fréquence des termes **TF** nous aide à identifier le nombre de fois que le terme ou le mot apparaît dans une ressource. Elle est représentée comme suit :

$$TF (M_i) = \frac{n_{Mi}}{N} \quad (3.1)$$

Où :

¹² <https://www.facebook.com/>

¹³ <https://fr.wikipedia.org/wiki/Facebook>

- n_{M_i} : désigne la fréquence du mot clé M_i dans une ressource.
- N : désigne le nombre total des mots clés dans une ressource.

IDF est calculé par la formule suivante :

$$IDF (M_i) = \log \left(\frac{R}{n_j} \right) \quad (3.2)$$

Où :

- R : désigne le nombre total de ressources de chaque utilisateur.
- n_j : désigne le nombre de ressources contenant le mot clé M_i .

Puis le poids **TF-IDF** pour le mot clé M_i dans la ressource est égale à :

$$Poids (M_i) = TF (M_i) \times IDF (M_i) \quad (3.3)$$

2.2.2. Extraction des données

Nous avons extrait les données à partir du réseau social Facebook, tel que : les informations personnelles, les recherches effectuées par des requêtes (mots clés) et leurs dates, les mots clés des ressources (image, vidéo) visités et leurs liens.

A cet effet, nous avons récupéré les données sur les utilisateurs non actifs qui ont effectué des recherches durant la période du Janvier 2021 au Septembre 2021, ainsi que les ressources consultées. La **figure 9** représente l'extraction des données à partir le réseau social Facebook.

ID Profil; Interets; Poids; Requete(mot clé); Date; Description (mots clés); Type ressource et leur lien						
1; Algo/Java; 0,5/ 0,7; Robe; 22 Juin 2021 13:00:00; Robe soirée Robe blanche femme ; Image	https://www.facebook.com/robosoireealcyceparis/photos/a.691713430905994/1731489356928391					
1; Algo/Java; 0,5/0,7; Robe; 24 Juin 2021 16:00:00; Location Robe blanche; Image	https://www.facebook.com/location.rob.es.blanches.blida/photos/a.104172854277909/525350542160136/					
1; Algo/Java; 0,5/ 0,7; Java; 01 Mai 2021 10:00:00; Javascript room room js; Image	https://www.facebook.com/roomjs/photos/a.125274034834784/545936816101835					
1; Algo/Java; 0,5/0,7; Java; 01 Mai 2021 11:00:00; Javascript room room js; Image	https://www.facebook.com/roomjs/photos/a.125274034834784/545936816101835					
1; Algo/Java; 0,5/0,7; Java; 01 Mai 2021 13:00:00; Learn Web Developement ; Image	https://www.facebook.com/roomjs/photos/a.125274034834784/544821716213345					
1; Algo/Java; 0,5/ 0,7; Java; 01 Mai 2021 14:00:00; Java developer class room; Image	https://www.facebook.com/photo/?fbid=2613074732171911&set=g.424586427906641					
1; Algo/Java; 0,5/ 0,7; Java; 01 Mai 2021 16:00:00; Microsoft azure Microsoft windows ; Image	https://www.facebook.com/photo/?fbid=4456118147778905&set=g.424586427906641					
1; Algo/Java; 0,5/ 0,7; Java; 02 Mai 2021 08:00:00; Microsoft azure Microsoft windows ; Image	https://www.facebook.com/photo/?fbid=4456118147778905&set=g.424586427906641					
1; Algo/Java; 0,5/ 0,7; Java; 02 Mai 2021 09:00:00; Microsoft azure Microsoft windows ; Image	https://www.facebook.com/photo/?fbid=4456118147778905&set=g.424586427906641					
1; Algo/Java; 0,5/ 0,7; Python; 05 Aout 2021 10:00:00; Méthodes importantes Python ; Image	https://www.facebook.com/photo/?fbid=1775992769267608&set=g.270834543757117					
1; Algo/Java; 0,5/ 0,7; Python; 05 Aout 2021 13:00:00; Méthodes importantes Python; Image	https://www.facebook.com/photo/?fbid=1775992785934273&set=g.270834543757117					
1; Algo/Java; 0,5/ 0,7; Python; 05 Aout 2021 14:00:00; Dimentionel reduction Pyhton; Image	https://www.facebook.com/photo/?fbid=3825436917561919&set=g.270834543757117					
1; Algo/Java; 0,5/ 0,7; Python; 06 Aout 2021 08:00:00 ; Python syllabus Cours methodes python; Image	https://www.facebook.com/photo/?fbid=181197664053353&set=g.270834543757117					
1; Algo/Java; 0,5/ 0,7; sandale; 05 Aout 2021 11:00:00 ; Sandale sandale claquette femme; Image	https://www.facebook.com/marketplace/109160299102441/?hoisted_items=251672783002688					
1; Algo/Java; 0,5/ 0,7; sandale; 20 Aout 2021 13:00:00 ; Sandale sandale claquette femme; Image	https://www.facebook.com/marketplace/109160299102441/?hoisted_items=251672783002688					

Figure 9 : Extraction des données à partir du réseau social Facebook

Cette base contient sept colonnes : ID de profil, liste des anciens centres d'intérêts, liste des poids de chaque intérêt, la recherche, la date et le temps de recherche, la description du ressource consultée (ensemble des mots clés) et le type de la ressource (image, vidéo) et son lien.

2.2.3. Prétraitement

Dans le but d'avoir des résultats fiables, nous avons procédé à la vérification des données en supprimant les données contiennent des erreurs (lien de ressource erroné où inexistant), des valeurs manquantes (ressource manque d'une description) ou d'autres types d'incohérence qui compliquent parfois l'analyse.

2.2.4. Jeu de données

Nous avons organisé les informations résultantes du prétraitement dans une base de données pour faciliter le processus de traitement et d'évolution, comme le montre le schéma conceptuel ci-dessus (Figure 10).

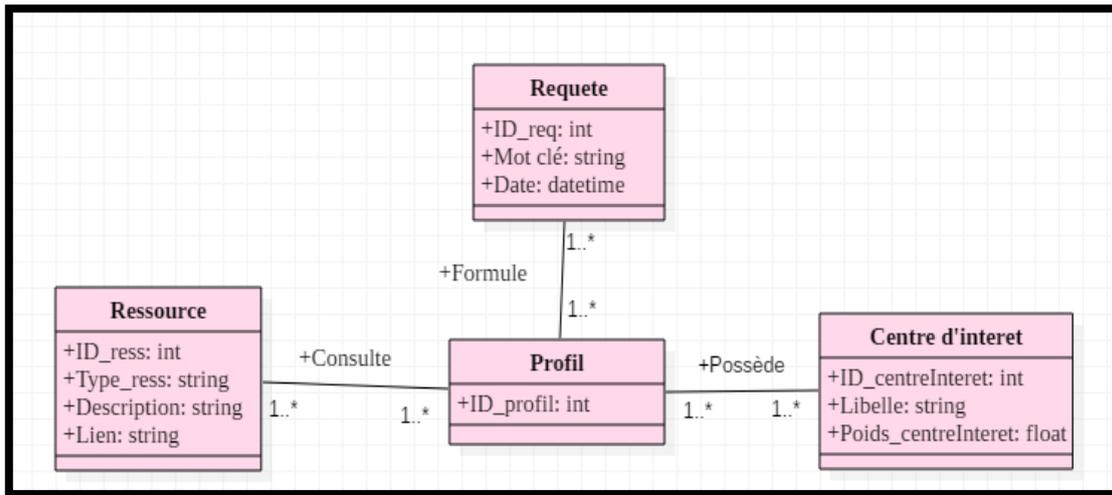


Figure 10 : Schéma conceptuel des informations résultantes

2.3. Espace utilisateur

Selon les types des utilisateurs mentionnés en chapitre 2 (*cf chap.2 §4.*), l'utilisateur non actif concerné par notre travail, n'interagit pas dans les réseaux sociaux (aucuns commentaires, aucuns partages, ...etc.) et ne possède aucun ami (très peu d'amis), les causes de son l'inactivité peuvent être : médicale, familiale, malicieuse, piratage, ...etc.

Un utilisateur utilise les réseaux sociaux pour des intérêts positifs, par exemple : informations médicaux, sports, cultures, ...etc. Pour les côtés négatifs, il peut être un compte malveillant ou faux profil pour voler les informations des autres utilisateurs ou pour le piratage. Nous distinguons les fonctionnalités de l'espace utilisateur dans le système :

- ❖ **Consultation** : chaque utilisateur peut consulter ses propres informations sociales : personnelles, ses centres d'intérêts, ainsi que les statistiques de ses recherches.
- ❖ **Mise à jour profil** : l'utilisateur peut modifier ses propres informations personnelles sociales.

2.4. Espace administrateur

L'administrateur est le responsable du système, il se charge de déclenchement de l'évolution des intérêts du profil social. Nous distinguons les fonctionnalités de l'administrateur suivantes :

- ❖ **Visualisation** : il peut consulter les informations personnelles, les centres d'intérêts de tous les utilisateurs.

- ❖ **Statistiques** : l'administrateur a le privilège de générer différentes statistiques produites par l'interaction des utilisateurs avec le système.
- ❖ **Evolution** : est le noyau de notre travail, elle vise à ne pas figer le profil social non actif, mais à la rendre plutôt une entité dynamique. A cet effet, nous nous intéressant à la détection des centres d'intérêts implicitement à partir des recherches des utilisateurs et les ressources consultées.

3. Evolution du profil social

Les intérêts sociaux de chaque utilisateur changent à travers le temps ce qui nécessite une évolution régulière et pertinente. Notre solution se devise en deux étapes, dans un premier temps nous avons effectué la pondération des requêtes (mots clés) pour détecter les mots clés pertinents les plus représentatives du profil social non actif. Vient par la suite l'étape d'enrichissement par l'ajout des nouveaux centres d'intérêts, nous nous sommes inspirés de la mesure de similarité. Elle consiste à détecter des profils similaires au profil concerné par l'évolution en calculant la similarité entre les centres d'intérêts du profil non actif et les profils existants dans la base de données. La **figure 11** montre le fonctionnement du processus évolution du profil social.

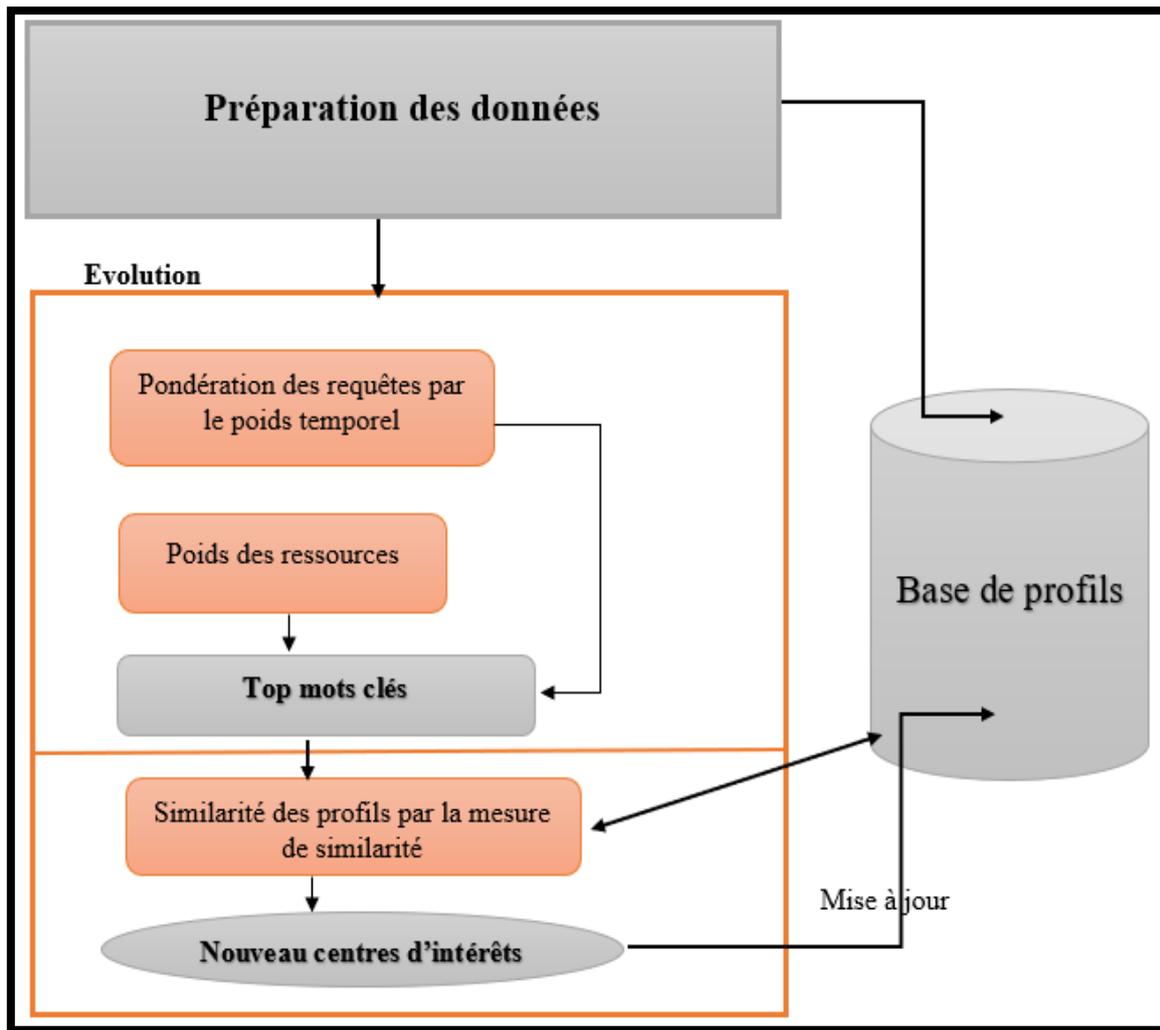


Figure 11 : Processus d'évolution du profil social

3.1. Pondération des requêtes

Pondération des requêtes se divise en deux étapes, dans un premier temps nous allons pondérer toutes les requêtes pour extraire les mots avec leurs poids. Vient par la suite la sélection des tops mots qui représenteront les centres d'intérêts de l'utilisateur (des nouveaux intérêts ou la mise à jour de ceux existant).

3.1.1. La pondération temporelle

L'utilisateur lance des requêtes/ recherches (contenant des mots clés), dans des moments et périodes du temps, pour obtenir l'information recherchée. Dans notre travail, nous nous sommes inspirés du travail [On-At, 2017] (cf chap.2 §7.2.) et nous considérons dans notre cas les requêtes comme des tags. Nous pondérons le poids de chaque mot clé avec une fonction temporelle linéaire inverse (est une technique de calcul du poids du mot clé de chaque

instance (période) selon la pertinence estimée en fonction du temps) qui va permettre de diminuer ou d'augmenter la pertinence des mots clés.

Les étapes de calcul du poids temporel d'un mot clé sont décrite comme suit :

1. Nous calculons la période entre le temps ancien et le temps actuel du mot clé M_i de chaque utilisateur par la formule :

$$Période (M_i) = |temps\ actuel - temps\ ancien| \quad (3.4)$$

2. Nous calculons le poids par période du mot clé M_i en utilisant la fonction temporelle linéaire inverse. Nous appliquons la formule (3.5) :

$$Poids_{période} (M_i) = \frac{1}{Période(M_i)+1} \quad (3.5)$$

3. Nous calculons le poids du mot clé en utilisant la somme des poids par période du mot par rapport à la somme de tous les poids par période des mots de chaque utilisateur. Ce calcul est donné par la formule :

$$Poids (M_i) = \frac{\sum Poids_{période}(M_i)}{\sum Poids_{période} (M_j)} \quad (3.6)$$

- M_i : représente le mot clé d'utilisateur.
- M_j : représente l'ensemble des tous les mots clés d'utilisateur.

3.1.2. Poids des mots clés

Après avoir calculé le poids de chaque mot clé dans l'étape précédente, nous améliorons ce poids par la prise en considération d'autres facteurs à savoir : l'existence du mot dans l'ancien ensemble des centres d'intérêts de l'utilisateur et dans les ressources consultées.

A cet effet, nous distinguons les cas suivants :

- a. **Cas 1** : Mot appartient à {requêtes & ressources & centres d'intérêts}, le poids est calculé comme suit :

$$Poids (M_i) = (Pm + Pr + Pc) / 3 \quad (3.7)$$

Où :

- Pm : désigne le poids des requêtes contenir le mot M_i de l'utilisateur.
- Pr : désigne le poids des ressources contenir le mot M_i de l'utilisateur.
- Pc : désigne le poids des centres d'intérêts contenir le mot M_i de l'utilisateur.

b. **Cas 2** : Mot appartient à {requêtes & ressources}, le poids est calculé comme suit :

$$Poids (M_i) = (Pm * 2 + \Sigma Pr) / (1 + \Sigma (Pm + Pr)) \quad (3.8)$$

c. **Cas 3** : Mot appartient {requêtes & centres d'intérêts}, le calcul est donné par cette formule (3.9) :

$$Poids (M_i) = (Pm + Pc) / 2 \quad (3.9)$$

d. **Cas 4** : Mot appartient {requêtes}, le poids est égal à celui déjà calculé de l'étape précédente.

e. **Cas 5** : Mot appartient {ressources & centres d'intérêts}, le poids est calculé par la formule suivante :

$$Poids (M_i) = (Pr + Pc) / 2 \quad (3.10)$$

f. **Cas 6** : Mot appartient {ressources}, nous appliquons cette formule :

$$Poids (M_i) = \frac{\Sigma Prc}{Nb R} \quad (3.11)$$

Où :

- **Prc** : représente le poids des mots communs entre les ressources d'utilisateur.
- **Nb R** : représente le nombre total des ressources par l'utilisateur.

La valeur du poids du mot clé est entre **0** et **1**, l'algorithme (**Algorithme 1**) montre le fonctionnement de calcul du poids d'un mot clé de chaque utilisateur.

Fonction calculer Poids (requête, utilisateur, temps actuel, temps ancien, liste mot clé) {

Variable Double poids ;

$P \leftarrow$ période entre temps actuel et temps ancien d'une requête ;

$P_p \leftarrow$ poids par période de chaque requête ; $P_t \leftarrow$ poids par période de tous les requêtes d'utilisateur ;

$P_c \leftarrow$ poids (centre intérêt) ; $P_m \leftarrow$ poids (requête) ; $P_r \leftarrow$ poids (ressource) ;

$N_b R \leftarrow$ nombre de ressources par l'utilisateur ; $P_{rc} \leftarrow$ poids mot commun entre ressources ;

Début

1 : Pour chaque requête d'un utilisateur

Faire

2 : $P = |\text{temps actuel} - \text{temps ancien}|$;

3 : $P_p = \frac{1}{P+1}$;

4 : $P_m = \frac{\sum P_p}{\sum P_t}$; **Fait** ;

5 : Pour $i \leftarrow 0$ à taille (liste mot clé) **Faire**

6 : Si liste mot clé [i] appartient à {requêtes et ressources et centres d'intérêts} **Alors**

7 : poids = $(P_c + P_r + P_m) / 3$; **Fin si** ;

8 : Sinon Si liste mot clé [i] appartient à {requêtes et ressources} **Alors**

9 : poids = $(P_m * 2 + \sum P_r) / (1 + \sum (P_m + P_r))$; ; **Fin si** ;

10 : Sinon Si liste mot clé [i] appartient à {requêtes et centres d'intérêts} **Alors**

11 : poids = $(P_m + P_c) / 2$; **Fin si** ;

12 : Sinon Si liste mot clé [i] appartient à {requêtes} **Alors**

13 : poids = P_m ; **Fin si** ;

14 : Sinon Si liste mot clé [i] appartient à {ressources et centres d'intérêts} **Alors**

15 : poids = $(P_r + P_c) / 2$; **Fin si** ;

16 : Sinon liste mot clé [i] appartient à {ressources} **Alors**

17 : poids = $\sum P_{rc} / N_b R$; **Fin si** ; **Fait** ;

18 : Retourner poids ;

19 : Fin ;

}

Algorithme 1 : Calcul des poids des mots de chaque utilisateur

3.1.3. Top Mots clés

Le poids permet le classement des mots selon leur importances et donc choisir les représentatifs pour l'utilisateur. A cet effet, nous procédons à la sélection des tops mots clés dont leurs poids dépassent le seuil qui est obtenu comme suit (3.12) :

$$Seuil = \frac{\sum Poids (M_i)}{\sum_{i=1}^n N_{M_i}} \quad (3.12)$$

Où :

- M_i : les mots clés appartient à {requêtes / ressources /centres d'intérêts} par l'utilisateur.
- N_{M_i} : le nombre des mots clés appartient à {requêtes / ressources /centres d'intérêts} par l'utilisateur.

Ces mots représentent les centres d'intérêts de l'utilisateur non actif. Cet ensemble des intérêts peut être pauvre en information vue l'inactivité de l'utilisateur, pour pallier ce problème nous proposons de l'enrichir par les centres d'intérêts des autres utilisateurs.

3.2. Mesure de similarité

Pour enrichir davantage les intérêts de l'utilisateur nous proposons d'ajouter d'autres mots/intérêts. Pour cela, les profils similaires au profil non actif peuvent être des sources de données. L'enrichissement consiste à calculer la similarité entre le profil non actif (ou profil candidat) et les profils existants (ou profils cibles) dans la base de données, plus la valeur de similarité est élevée, plus les centres d'intérêts du profil cible sont susceptibles d'être intéressants pour le profil candidat.

Il existe de nombreuses formules mathématiques qui peuvent être utilisées pour calculer la similarité entre les utilisateurs, dans notre travail nous avons choisi d'utiliser la fonction de similarité décrite dans [Elachkar et al., 2020], nous avons un ensemble des profils cibles P_i , où chaque profil contient un ensemble des centres d'intérêts C_i et leurs poids W_i . Nous avons aussi un profil candidat P_j contient des centres d'intérêts C_j et leurs poids W_j .

- ❖ Tout d'abord, nous construirons une matrice où les lignes représentent les profils P_i et le profil P_j et les colonnes représentent les centres d'intérêts C_i et C_j .

Si C_i appartient à P_i et C_j appartient à P_j , on mettra la valeur **1**. Sinon la valeur **0**.

- ❖ Ensuite, nous calculons la similarité S entre le profil candidat et le profil cible par la formule suivante :

$$S (P_i, P_j) = \frac{\sum N_{C_i \cap C_j}}{\sum N_{C_j}} \quad (3.13)$$

Où :

- $N_{C_i \cap C_j}$: le nombre de centre d'intérêts communs entre le profil cible P_i et le profil candidat P_j .
 - N_{C_j} : Le nombre de centres d'intérêts du profil candidat P_j .
- ❖ A la fin nous prenons la valeur de la similarité élevée entre le profil candidat et les autres profils cibles et nous allons enrichir le profil candidat par les centres d'intérêts non communs avec le profil cible le plus similaire. L'algorithme (**Algorithme 2**) ci-dessus réalise la fonction de similarité choisie.

Fonction Similarité (profil, liste profil []) {

Variable J ← profil candidat ; I ← profil cible ;

S ← similarité entre profil candidat et profil cible ;

M [] [] ← matrice contient des valeurs des profils cibles et du profil candidat ;

Ci ← Centre d'intérêts du profil cible ; Cj ← Centre d'intérêts du profil candidat ;

Nb ← nombre de centre d'intérêts communs entre profil cible et profil candidat ;

Début

1 : Pour chaque profil cible **Faire**

2 : Pour chaque profil candidat **Faire**

3 : Si Ci appartient à I **Alors**

4 : M [] [] = 1 ;

5 : Si Cj appartient à J **Alors**

6 : M [] [] = 1 ;

7 : Sinon

8 : M [] [] = 0 ; **Fin si ; Fin si ; Fait ; Fait ;**

9 : $S = \frac{\sum Nb}{\sum Cj}$;

10 : Prendre le max (S) ;

11 : Si Ci diffère à Cj **Alors**

12 : Cj ← Ci ; Ajouter centres d'intérêts Cj au profil J ; **Finsi ;**

13 : Retourner J ;

14 : Fin ;}

Algorithme 2 : Fonction de similarité

3. Diagrammes UML représentatifs du système

UML (Unified Modeling Language) est une méthode de modélisation orientée objet développée dans le but de définir la notation standard pour la modélisation des applications, construites à l'aide d'objets.

Relativement à notre système se base sur les trois modèles :

- ❖ Le diagramme des cas d'utilisation.
- ❖ Le diagramme de séquence.

❖ Le diagramme de classes.

4.1. Diagramme des cas d'utilisation

Le diagramme des cas d'utilisation est un diagramme UML utilisé pour une représentation du comportement fonctionnel d'un système logiciel. Dans un diagramme des cas d'utilisation, un acteur représente un rôle joué par une personne qui interagit avec le système. La **figure 12** ci-dessus représente le diagramme des cas d'utilisation de notre système.

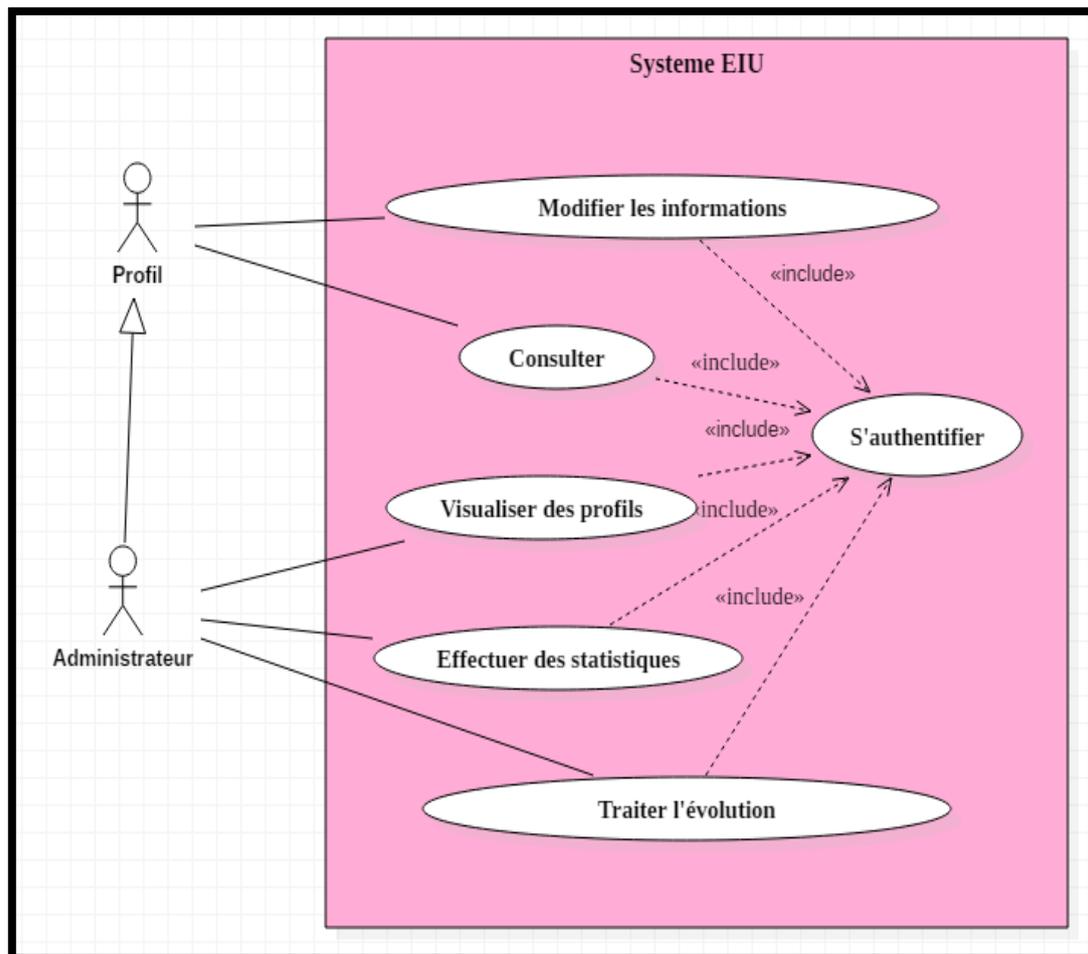


Figure 12 : Diagramme des cas d'utilisation

Description de diagramme des cas utilisation

Un acteur est une entité externe qui interagit avec le système, dans notre travail nous avons les acteurs : administrateur et profil de chaque utilisateur. Ce **tableau 1** ci-dessus présente les acteurs et leurs rôles.

Acteur	Rôles
Administrateur	<p>Authentification : administrateur doit saisir son pseudo et le mot de passe correct pour accéder à l'espace administrateur.</p> <p>Traitement d'évolution : après authentification, l'administrateur est le responsable du processus d'évolution du profil social.</p> <p>Visualisation : il consulte toutes les informations personnelles sociales et les centres d'intérêts des profils.</p> <p>Statistiques : consulter les résultats en forme des statistiques sur l'interaction de l'utilisateur avec le système.</p>
Profil	<p>Authentification : chaque profil doit saisir son pseudo et son mot de passe correct pour accéder à l'espace utilisateur.</p> <p>Consultation : il peut consulter ses propres informations sociales : personnelles, ses centres d'intérêts et les statistiques des requêtes de recherche.</p> <p>Modification ses informations : il est la possibilité de changer ses propres informations personnelles sociales.</p>

Tableau 1 : Description des rôles des acteurs

4.2. Diagramme de séquence

Le diagramme de séquence est une modélisation en langage UML, une interaction entre les acteurs et le système par des messages pour exercer une fonction avant la fin de la ligne de vie.

4.2.1. Diagramme de séquence du cas « S'authentifier »

L'utilisateur (administrateur ou profil) doit saisir le pseudo et le mot de passe correct pour accéder à son espace dans le système (**Figure 13**).

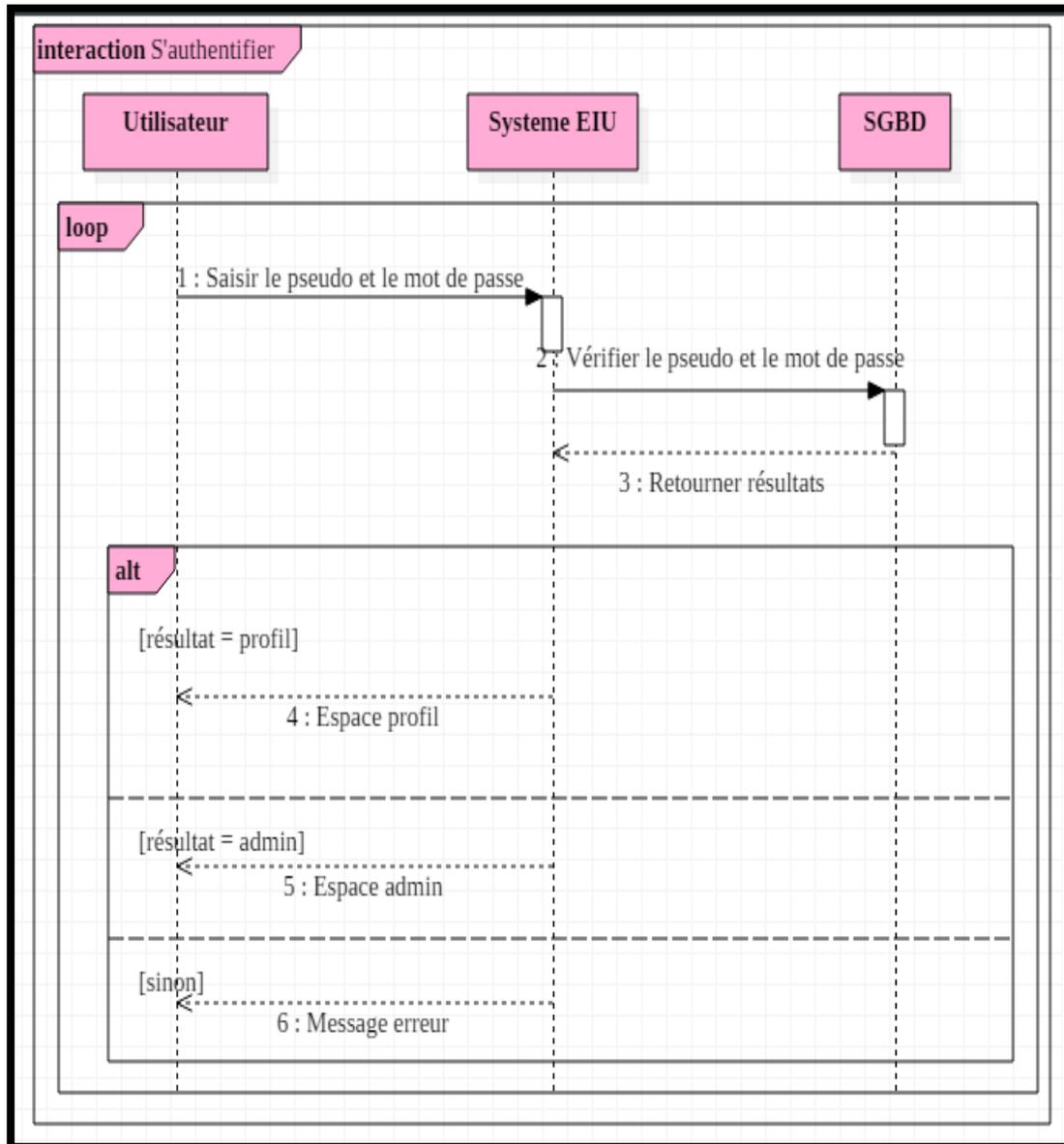


Figure 13 : Diagramme de séquence du cas « S'authentifier »

4.2.2. Diagramme de séquence du cas « Traiter l'évolution »

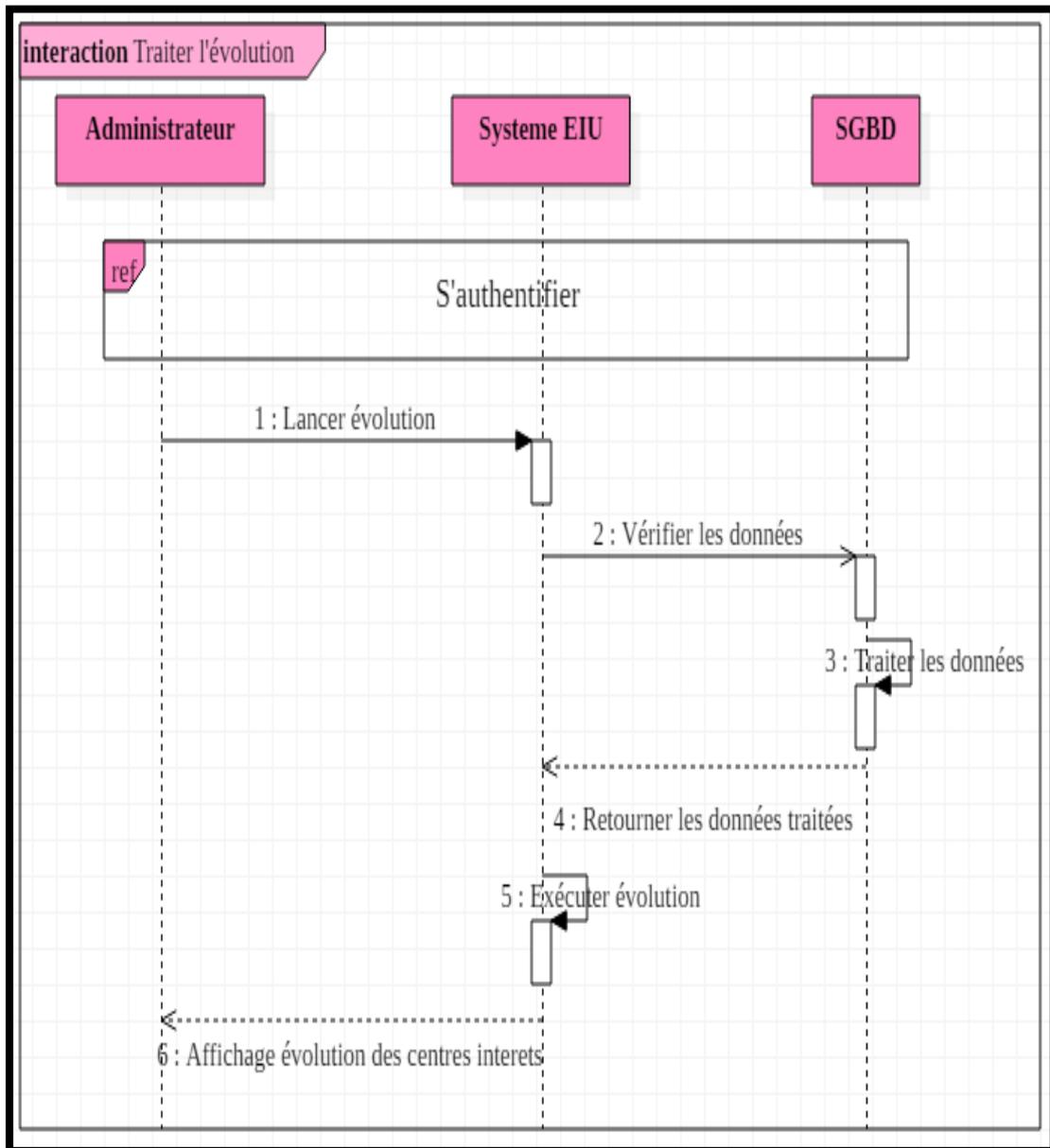


Figure 14 : Diagramme de séquence du cas « Traiter l'évolution »

4.3. Diagramme de classes

Le diagramme de classes décrit la structure interne du notre système, il identifie la structure des classes d'un système, y compris les propriétés et les méthodes de chaque classe. La **figure 15** représente notre diagramme de classes.

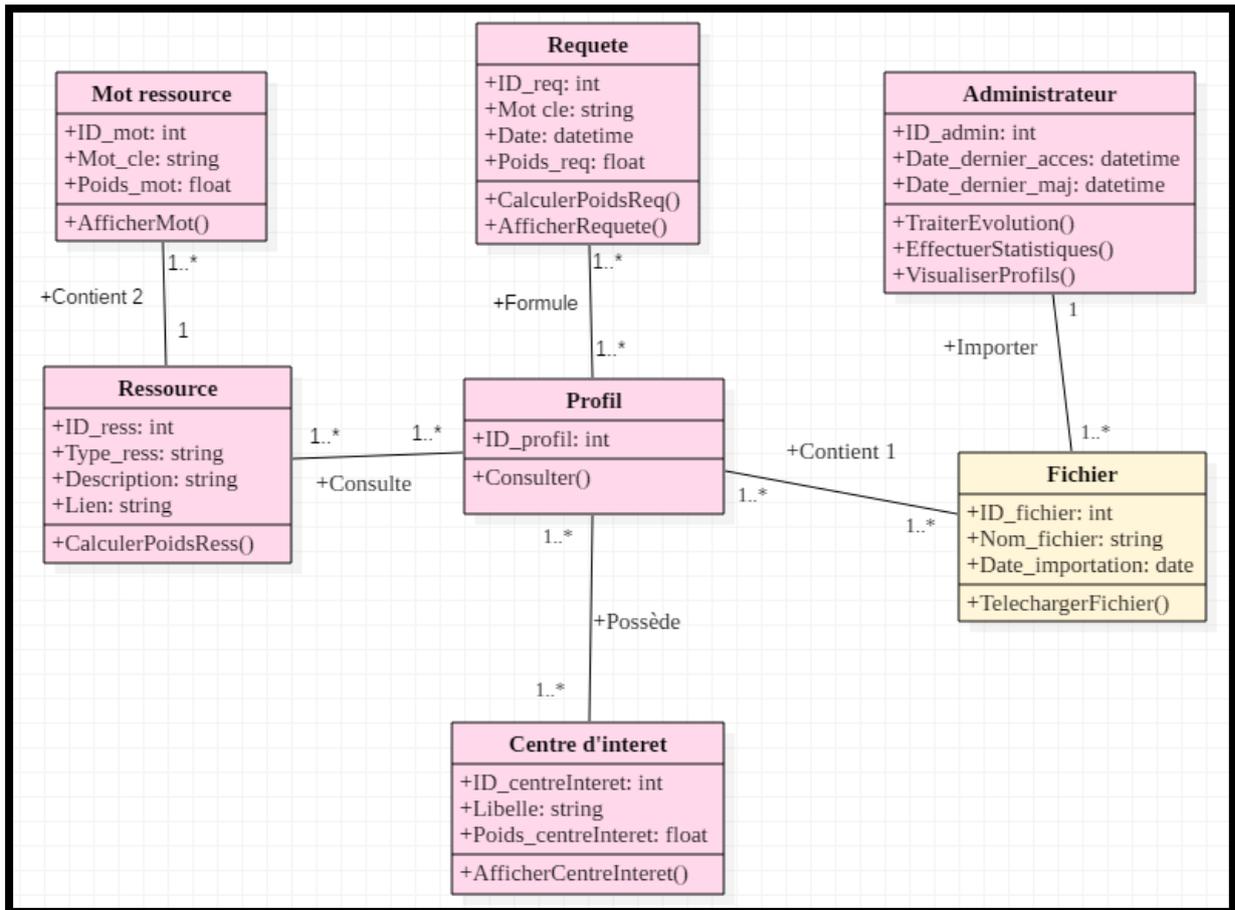


Figure 15 : Diagramme de classes

4.4. Passage du diagramme d'objet vers modèle relationnel

Pour réaliser la traduction du diagramme de classe en un modèle relationnel, nous appliquons un certain nombre de règles de passages afin d'obtenir le schéma suivant :

Administrateur (ID_admin, Date_dernier_acces, Date_dernier_maj)

Fichier (ID_fichier, Nom_fichier, Date_importation, *ID_admin)

Profil (ID_profil, *ID_fichier)

Centre d'intérêt (ID_centreInteret, Libelle, Poids_CentreInteret)

Requête (ID_req, Mot_cle, Date, Poids_req)

Ressource (ID_ress, Type_ress, Description, Lien)

Mot ressource (ID_mot, Mot_cle, Poids_mot, *ID_ress)

Possède (*ID centreInteret, *ID profil)

Formule (*ID req, *ID profil)

Consulte (*ID ress, *ID profil)

5.Conclusion

Dans ce chapitre, nous avons présenté l'architecture globale de notre système et sa description. Ensuite, nous avons détaillé notre solution pour l'évolution des centres d'intérêts des utilisateurs non actifs. Nous avons commencé par la construction de l'ensemble des mots et leurs poids pour constituer la liste des centres d'intérêts. Vu l'inactivité de l'utilisateur, l'ensemble des intérêts est pauvre en mots, pour cela nous avons proposé de l'enrichir avec les intérêts des autres utilisateurs en utilisant une fonction de similarité. Par la suite, nous avons modélisé notre système à l'aide des outils du langage UML. Tous les éléments sont désormais mis en place pour entamer la mise en œuvre de notre système, cela fera l'objet du prochain chapitre.

Chapitre 4

*Mise en œuvre et
expérimentation du système
d'évolution des profils sociaux*

1. Introduction

Après l'étude et la conception de notre système, nous arrivons dans ce chapitre à la mise en œuvre de notre application desktop, qui va nous permettre d'arriver aux objectifs fixés précédemment. Dans la première partie, nous allons présenter les outils utilisés pour l'implémentation et dans la deuxième partie, nous allons présenter les principales interfaces qui la composent à travers des fenêtres de capture.

2. Environnement de développement

Nous décrivons les outils d'environnement de développement choisis, pour réaliser notre système (**Figure 16**) :

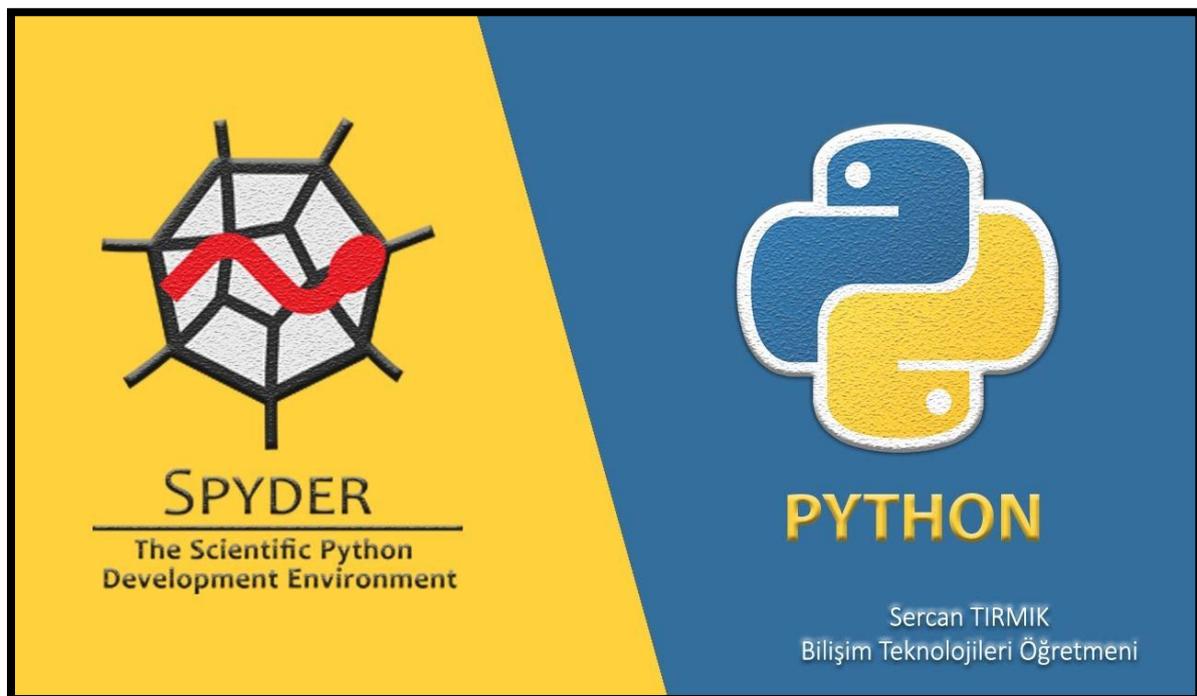


Figure 16 : Les outils d'environnement de développement

2.1. Python

Python¹⁴ est un langage de programmation interprété, multiparadigme et multiplateformes. Il favorise la programmation impérative structurée, fonctionnelle et orientée objet¹⁵. Nous avons utilisé le langage python pour des raisons comme suit :

¹⁴ <https://www.python.org/downloads/>

¹⁵ [https://fr.wikipedia.org/wiki/Python_\(langage\)](https://fr.wikipedia.org/wiki/Python_(langage))

Chapitre 4 Mise en œuvre et expérimentation du système d'évolution des profils sociaux

- ❖ Syntaxe très simple et combiné à des types de données évoluées (listes, dictionnaires, ...etc.).
- ❖ Plus court 5 à 3 fois qu'un programme en java ou en C++, ce qui implique un temps de développement 5 à 10 fois plus court.

2.2. Spyder

Spyder¹⁶ est un environnement de développement pour Python, libre et multiplateforme. En comparaison avec d'autres IDE pour le développement scientifique, il a un ensemble unique de fonctionnalités - multiplateforme, open-source, écrit en Python et disponible sous une licence non-copyleft. Il est extensible avec des plugins, comprend le support d'outils interactifs pour l'inspection des données et incorpore des instruments d'assurance de la qualité et d'introspection spécifiques au code Python¹⁷.

2.3. PyQt

PyQt est un module libre qui permet de lier le langage Python avec la bibliothèque Qt distribué sous deux licences : une commerciale et la GNU GPL. Il permet ainsi de créer des interfaces graphiques en Python. Une extension de Qt Creator (utilitaire graphique de création d'interfaces Qt) permet de générer le code Python d'interfaces graphiques. Les applications utilisant PyQt : Spyder, Calibre, ...etc¹⁸.

3. Interfaces de l'application

Dans cette partie, nous présentons notre application desktop et les différentes interfaces homme/machine qui permet de faciliter l'utilisation du système.

3.1. Authentification

L'utilisateur ou administrateur doit saisir le pseudo et mot de passe correct pour accéder à son espace (**Figure 17**).

¹⁶ <https://www.anaconda.com/products/individual>

¹⁷ [https://fr.wikipedia.org/wiki/Spyder_\(logiciel\)](https://fr.wikipedia.org/wiki/Spyder_(logiciel))

¹⁸ <https://fr.wikipedia.org/wiki/PyQt>

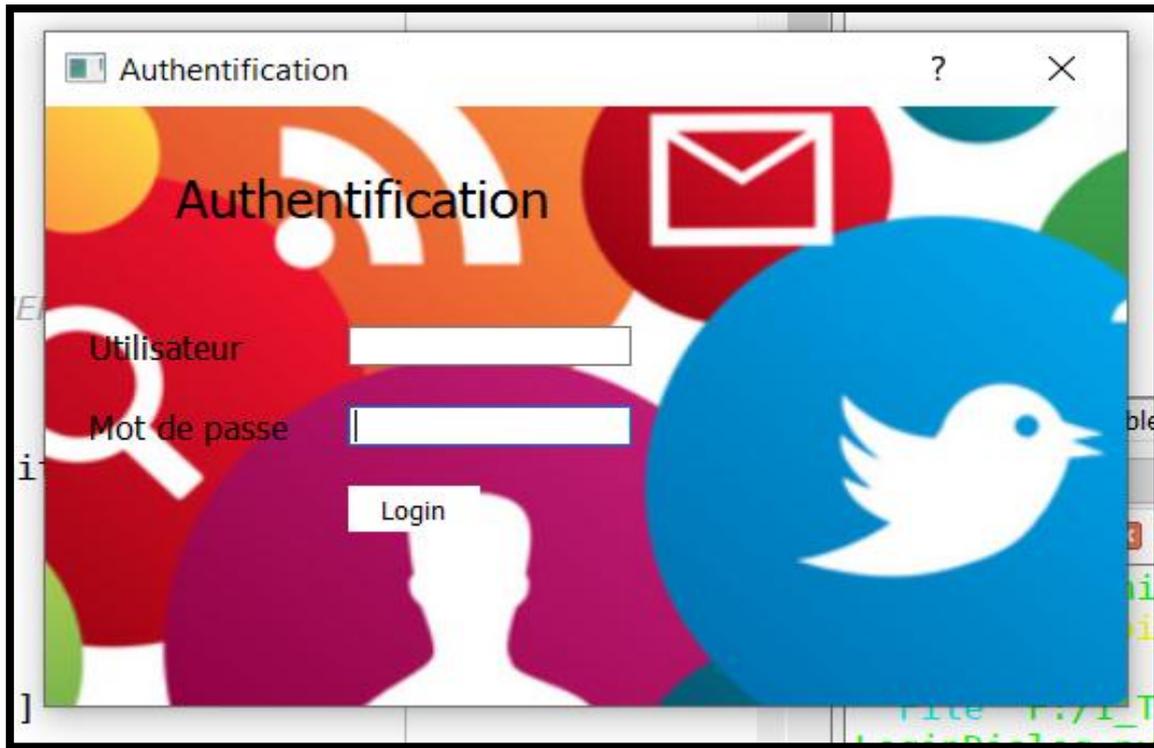


Figure 17 : Page authentification

3.2. Espace utilisateur

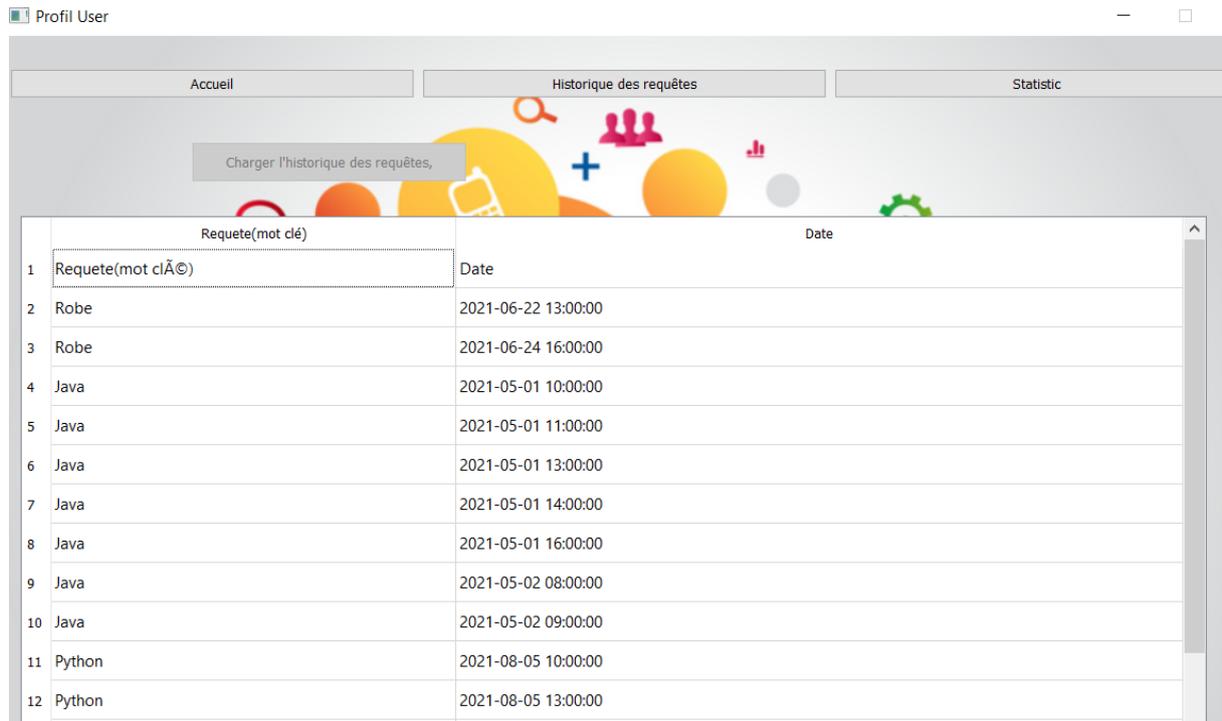
Après l'authentification, l'utilisateur accède à son espace personnel où il pourra interagir avec le système en utilisant des différentes fonctionnalités :



Figure 18 : Espace utilisateur

3.2.1. Consultation des requêtes

L'utilisateur a la possibilité de consulter l'historique de ses recherches (**Figure 19**).



	Requete(mot clé)	Date
1	Requete(mot clé)	Date
2	Robe	2021-06-22 13:00:00
3	Robe	2021-06-24 16:00:00
4	Java	2021-05-01 10:00:00
5	Java	2021-05-01 11:00:00
6	Java	2021-05-01 13:00:00
7	Java	2021-05-01 14:00:00
8	Java	2021-05-01 16:00:00
9	Java	2021-05-02 08:00:00
10	Java	2021-05-02 09:00:00
11	Python	2021-08-05 10:00:00
12	Python	2021-08-05 13:00:00

Figure 19 : Consultation l'historique des recherches

3.2.2. Consultations statistiques

L'utilisateur a la possibilité de consulter ses centres d'intérêts sous forme des statistiques (**Figure 20**).



Figure 20 : Consultation statistiques des centres d'intérêts

3.3. Espace administrateur

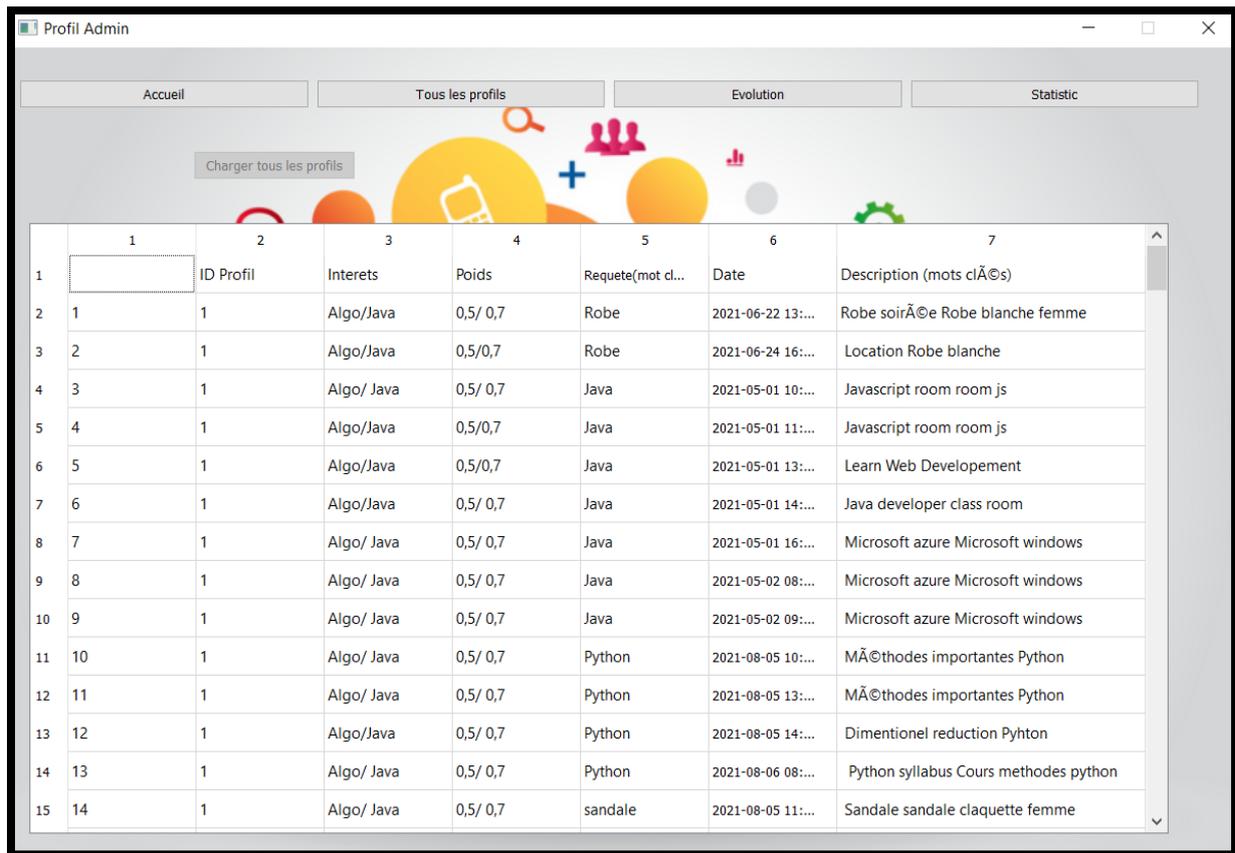
Un espace est consacré à l'administrateur, à travers lequel il pourra, d'une part, gérer l'évolution des centres d'intérêts des profils utilisateurs non actifs d'autre part, avoir une vue globale sur les profils. La **figure 21** représente l'espace d'accueil de l'administrateur.



Figure 21 : Espace administrateur

3.3.1. Visualisation des profils

L'administrateur peut consulter les informations personnelles, les requêtes de recherche et leurs dates, ainsi les ressources consultées de tous les utilisateurs (**Figure 22**).



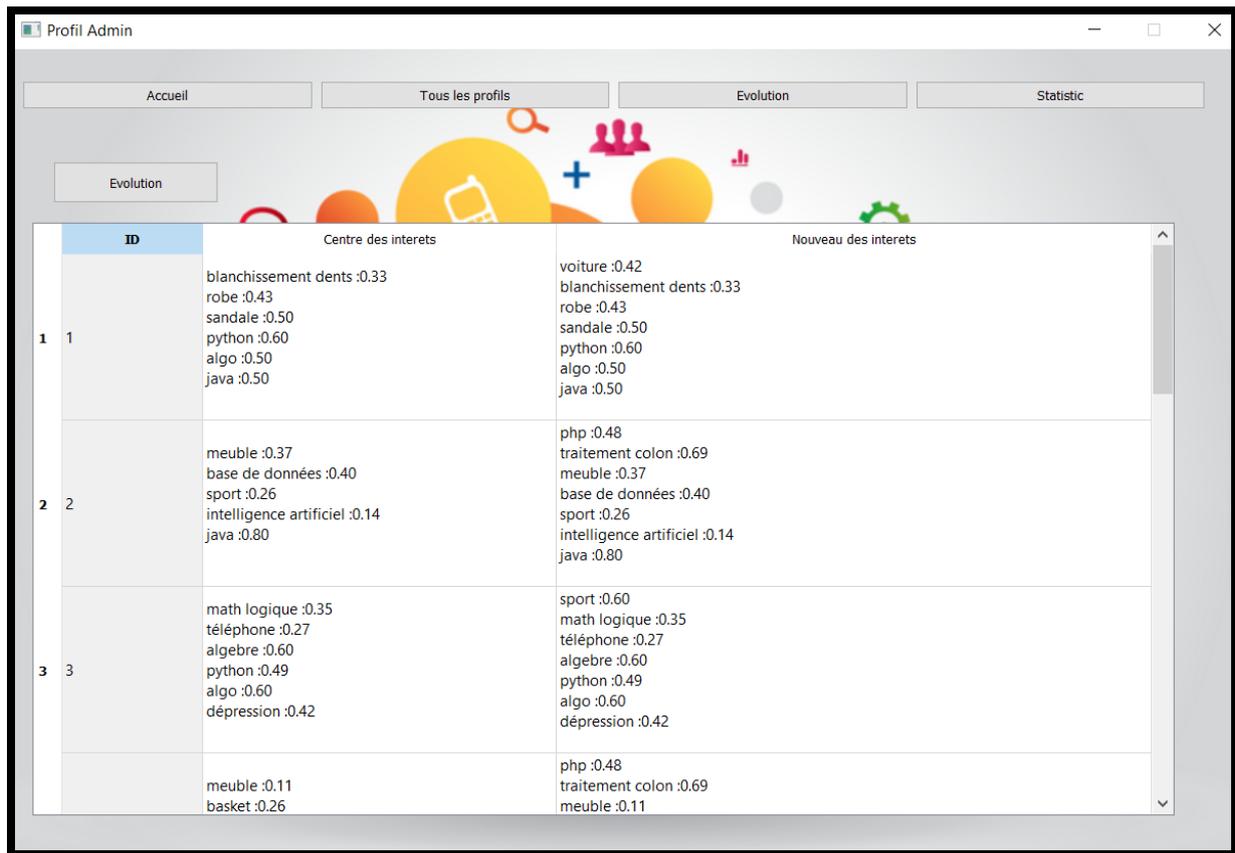
The screenshot shows a web application window titled 'Profil Admin'. At the top, there are four navigation tabs: 'Accueil', 'Tous les profils', 'Evolution', and 'Statistic'. Below the tabs is a decorative header with colorful icons (magnifying glass, people, plus, bar chart, gear) and a button labeled 'Charger tous les profils'. The main content is a table with 7 columns and 15 rows. The columns are: 1 (Index), 2 (ID Profil), 3 (Interets), 4 (Poids), 5 (Requete(mot cl...), 6 (Date), and 7 (Description (mots clé(s)).

1	2	3	4	5	6	7	
1		ID Profil	Interets	Poids	Requete(mot cl...	Date	Description (mots clé(s)
2	1	1	Algo/Java	0,5/ 0,7	Robe	2021-06-22 13:...	Robe soirÃ©e Robe blanche femme
3	2	1	Algo/Java	0,5/0,7	Robe	2021-06-24 16:...	Location Robe blanche
4	3	1	Algo/ Java	0,5/ 0,7	Java	2021-05-01 10:...	Javascript room room js
5	4	1	Algo/Java	0,5/0,7	Java	2021-05-01 11:...	Javascript room room js
6	5	1	Algo/Java	0,5/0,7	Java	2021-05-01 13:...	Learn Web Developement
7	6	1	Algo/Java	0,5/ 0,7	Java	2021-05-01 14:...	Java developer class room
8	7	1	Algo/ Java	0,5/ 0,7	Java	2021-05-01 16:...	Microsoft azure Microsoft windows
9	8	1	Algo/ Java	0,5/ 0,7	Java	2021-05-02 08:...	Microsoft azure Microsoft windows
10	9	1	Algo/ Java	0,5/ 0,7	Java	2021-05-02 09:...	Microsoft azure Microsoft windows
11	10	1	Algo/ Java	0,5/ 0,7	Python	2021-08-05 10:...	MÃ©thodes importantes Python
12	11	1	Algo/ Java	0,5/ 0,7	Python	2021-08-05 13:...	MÃ©thodes importantes Python
13	12	1	Algo/Java	0,5/ 0,7	Python	2021-08-05 14:...	Dimensionel reduction Pyhton
14	13	1	Algo/ Java	0,5/ 0,7	Python	2021-08-06 08:...	Python syllabus Cours methodes python
15	14	1	Algo/ Java	0,5/ 0,7	sandale	2021-08-05 11:...	Sandale sandale claquette femme

Figure 22 : Visualisation des profils

3.3.2. Evolution des profils sociaux

L'administrateur est le responsable du processus d'évolution des profils utilisateurs, il possède la capacité de lancer manuellement le processus d'évolution du profil social (Figure 23).



ID	Centre des interets	Nouveau des interets
1	blanchissement dents :0.33 robe :0.43 sandale :0.50 python :0.60 algo :0.50 java :0.50	voiture :0.42 blanchissement dents :0.33 robe :0.43 sandale :0.50 python :0.60 algo :0.50 java :0.50
2	meuble :0.37 base de données :0.40 sport :0.26 intelligence artificiel :0.14 java :0.80	php :0.48 traitement colon :0.69 meuble :0.37 base de données :0.40 sport :0.26 intelligence artificiel :0.14 java :0.80
3	math logique :0.35 téléphone :0.27 algèbre :0.60 python :0.49 algo :0.60 dépression :0.42	sport :0.60 math logique :0.35 téléphone :0.27 algèbre :0.60 python :0.49 algo :0.60 dépression :0.42
	meuble :0.11 basket :0.26	php :0.48 traitement colon :0.69 meuble :0.11

Figure 23 : Evolution des profils sociaux

3.3.3. Statistiques

L'une des principales fonctionnalités de notre application est de permettre à l'administrateur de consulter le comportement des utilisateurs avec le système par des statistiques sous forme graphique (**Figure 24**).



Figure 24 : Statistiques des poids des centres d'intérêts

4. Conclusion

Nous avons abordé à travers ce chapitre l'implémentation et la mise en œuvre de notre système. Tout d'abord, nous avons présenté les principaux outils qui nous ont permis de développer notre application. Ensuite, nous avons décrit les principales fonctionnalités de notre application, illustrant sa convivialité à travers les différentes interfaces qui la représentent, tout en détaillant la fonctionnalité la plus importante l'évolution des centres d'intérêts des utilisateurs non actifs.

Conclusion Générale

Conclusion générale

Les réseaux sociaux sont des sites Internet qui permettent aux utilisateurs, professionnels ou particuliers, de partager des informations. Chaque utilisateur doit créer un profil pour publier et consulter différents contenus : textes, photos, vidéos, ...etc. Ce sont des grands espaces de partage qui offrent la possibilité à des millions de personnes d'être interconnectées, indépendamment de leur situation géographique. Les intérêts de l'utilisateur évoluant au fil du temps dans la vie réelle, il en est de même pour ceux extraits depuis son réseau social : pertinents à un moment donné, ils peuvent ne plus être significatifs ultérieurement. Partant de ce constat, les principales informations qui permettent la détection des changements des centres d'intérêts ne sont pas ciblées sur l'utilisateur lui-même mais sur les éléments de son réseau social (liens entre les membres, informations qui circulent entre eux), l'évolution du profil social de l'utilisateur est donc liée à l'évolution de son réseau social.

Dans notre travail, nous nous sommes intéressés à l'évolution des intérêts sociaux des utilisateurs non actifs, cette évolution est basée sur l'analyse des différentes informations recherchées et des ressources consultées par les utilisateurs. A cet effet, nous avons proposé l'utilisation du poids temporel des informations recherchées (requêtes) et le poids des ressources consultées afin de trouver les tops mots clés représentant ainsi les nouveaux centres d'intérêts de chaque utilisateur. Par la suite nous avons proposé d'enrichir ces derniers ont les comparant avec ceux d'autres profils par l'application d'une mesure de similarité.

Bien que nous ayons réalisé les principaux objectifs définis, cependant nous envisageons quelques perspectives qui permettent l'amélioration et la conformité de notre travail entre autres :

- ❖ Etendre de notre système par l'ajout des groupes et des communautés.
- ❖ Etudier la possibilité d'appliquer TF-IDF sur les tags d'une ressource, lorsque ces derniers manquent de description.
- ❖ Prendre en considération le poids durant la similarité entre les profils.
- ❖ Comparer notre solution avec des solutions à base des algorithmes évolutifs (génétique, réseau de neurones, ...etc.).
- ❖ Etendre de notre solution pour permettre l'évolution et la mise à jour des centres d'intérêts de l'ensemble des profils non actifs.

Bibliographie

- [Amer-Yahia et al., 2007]** S. Amer-Yahia, M.Benedikt, P.Bouhannon, « Challenges in searching online communities », IEEE Data Eng., 2007.
- [Astrain et al., 2010]** J.J.Astrain, A.Cordoba, F.Echarte, J.Villadangos, « An algorithm for the improvement of tag-based social interest discovery », In Semapro 10: Proceedings of the fourth international conference on advances in semantic processing, Spain, 2010.
- [Basu et al., 1998]** C.Basu, H.Hirsh, W.Cohen, « Recommendation as classification : using social and content-based information in recommendation », Proceedings of the fifteenth national/tenth conference on Artificial intelligence/Innovative applications of artificial intelligence, American Association for Artificial Intelligence, USA, 1998.
- [Benammar et al., 2002]** A.Benammar, G.Hubert, J.Mothe, « Automatic profile reformulation using a local document analysis », European Conference on Information Retrieval, UK, 2002.
- [Berlingerio et al., 2013]** M.Berlingerio, M.Coscia, F.Giannotti, A.Monreale, D.Pedshrechi, « Multidimensional networks: foundation of structural analysis », World Wide Web, Springer Science Business Media, 2013.
- [Bilal et al., 2019]** M.Bilal, A.Gani, M.Ikram allah Lali, M.Marjani, N.Malik « Social Profiling: A Review, Taxonomy, and Challenges », Cyber psychology, Behavior, and Social Networking, vol.22, 2019.
- [Bilge et al., 2009]** L.Bilge, T.Strufe, D.Belzarotti, E.Kirda, « All your contacts are belong to us: automated identity theft attacks on social networks », In: Proceedings of the 18th international conference on World Wide Web, Madrid, Spain, 2009.
- [Boccaletti et al., 2006]** S.Boccaletti, V.Latora, Y.Moreno, D.Hwang, « Complex networks: Structure and dynamics », Journal Physics Reports, 2006.
- [Boulkrinat et al., 2019]** N.Boulkrinat, A.Meziane, N.Benblidia « Query-based Profile Evolution using Genetic Algorithm », 2019 IEEE/ACS 16th International Conference Of Computar Systems and Applications (AICCSA), Abu Dhabi, UAE, 2019.

- [Boyd et al., 2007]** D.Boyd, N.Ellison, « Social Network Sites: definition, history, and scholarship », Journal of computer –Mediated Communication, 2007.
- [Cabanac, 2011]** G.Cabanac, « Accuracy of inter-researcher Similarity Measures Based on Topical and Social Clues», An International Journal for all Quantitative Aspects of the Science of Science, Communication in science and Science policy, 2011.
- [Canut et al., 2015]** C.M.Canut, S.On-at, A.Péninou, F.Sédes, « Enrichissement du profil utilisateur à partir de son réseau social dans un contexte dynamique : application d’une méthode de pondération temporelle », International conférence In :33eme congrès Informatique des organisations et Systèmes d’Information et de Décision, France, 2015.
- [Cavazza, 2009]** F.Cavazza, « Une définition des médias sociaux », Usage numérique et transformation dégitale, doi://fredcavazza.net/2009/06/29/une-definition-des-medias-sociaux/, 2009.
<https://fredcavazza.net/2009/06/29/une-definition-des-medias-sociaux/> [Consulté le :28/05/2021]
- [Chen et al., 2002]** C.Chen, M.Chen, Y.Sun, « Pva: A self-adaptive personal view agent”, Journal of Intelligent Information Systems, 2002.
- [Chen et al., 2014]** C.Chen, D.Guan, Q.Su, « Feature set identification for detecting suspicious URLs using Bayesian classification in social networks », Journal Information Sciences, 2014.
- [Cheng et al., 2008]** Y.Cheng, G.Qiu, J.Bu, K.Liu, Y.Han, C.Wang, C.Chen, « Model Bloggers’ Interests Based on Forgetting Mechanism », In Proceedings of the 17th International Conference on World Wide Web, New York, USA, 2008.
- [Crabtree et al., 1998]** B.Crabtree, S.Soltysiak, M.Pp, I.Re, « Identifying and tracking changing interests », International Journal on Digital Libraries 2, 1998.
- [Cresci et al., 2015]** S.Cresci, R.Di Pietro, M.Petrocchi, A.Spognardi, M.Tesconi, « Fame for sale: Efficient detection of fake Twitter followers », Journal Decision Support System, 2015.
- [Cross et al., 2004]** R.Croos, A.Parker, « The hidden power of social networks: Understanding How Work really gets done in organizations », Edition Harvard Business Press, 2004.

- [Egele et al., 2015]** M.Egele, G.Stringhini, C.Kruegel, G.Vigna, « Towards Detecting Compromised Accounts on Social Networks », IEEE Transaction on Dependable and Secure Computing, 2015.
- [Ekstrand, et al., 2011]** M.D.Ekstrand, J.T.Riedl, J.A.Konstan, « Collaborative Filtering Recommender Systems », Found. Trends Hum.-Comput, Interact, 2011.
- [Elachkar et al., 2020]** I.Elachkar, H.Ouzif, H.Labridji, « Structural similarity measure of users profiles based on a weighted bipartite graphs », International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences 5th International Conference on Smart City Applications, Safranbolu, Turquie, 2020.
- [Ezpeleta et al., 2015]** E.Ezpeleta, U. Zurutuza, J.M. Gomez Hidalgo, « An Analysis of the Effectiveness of Personalized Spam Using Online Social Network Public Information », In A. Herrero, B.Baruque, J.Sedano, H.Quintian E.Corchado (Eds.), International Joint Conference: Cisis'15 and Iceute'15 (Vol. 369), Seville, Spain, 2015.
- [Fayon, 2008]** D.Fayon, « Web 2.0 et au-delà, nouveaux internautes : du surfeur à l'acteur », Ed Economia, 2008.
- [Fire et al., 2014]** M.Fire, D. Kagan, A. Elyashar, Y. Elovici, « Friend or foe? Fake profile identification in online social networks », Journal Social Network Analysis and Mining, 2014.
- [Fogues et al., 2015]** R.Fogues, M.Such, A.Espinosa, A.Garcia-Fornes, « Open Challenges in Relationship-Based Privacy Mechanisms for Social Network Services », International Journal of Human-Computer Interaction, vol.31, 2015.
- [Fondeur et al., 2006]** Y.Fondeur, F.Lhermite, « Social network sites: definitions, history, and scholarship », Journal of computer-mediated communication, 2006.
- [Freeman, 1978]** L.C.Freeman, « Centrality in social networks conceptual clarification », Journal Social networks, 1978.
- [Gauch et al., 2007]** S.Gauch, M.Speretta, A.Chandramouli, A.Micarelli, « User Profiles for Personalized Information Access », BRUSILOVSKY P., KOBSA A., NEJDL W. (dirs.), The Adaptive Web, Springer Berlin Heidelberg (Lecture Notes in Computer Science), 2007.

- [Girard, 2012]** A.Girard, , « L'intégration des médias sociaux dans les stratégies d'e-GRH : le cas du recrutement », thèse de doctorat, Université Montpellier 2, France, 2012.
- [Godoy et al., 2008]** D.Godoy, A.Amandi, « Hybrid content and tag-based profiles for recommendation in collaborative tagging systems », In Latin American web conference, Brazil, 2008.
- [Gupta, 2019]** P.Gupta, « Neural Information Extraction From Natural Language Text », thèse de doctorat, Faculté of Mathématiques, Computer Science and Statistics, Munich, Allemagne, 2019.
- [Guy et al., 2010]** I.Guy, N.Zwerdling, I.Ronen, D.Carmel, U.Eziel, « Social media recommendation based on people and tags », In Proceedings of the 33rd international ACM SIGIR conference on research and development in information retrieval, New York, USA, 2010.
- [Holland, 1992]** J.H.Holland, « Adaptation in natural and artificial and systems", Edition MIT Press, 1992.
- [Houmadi, 2007]** B. Houmadi, « Étude Exploratoire D'outils Pour Le Data Mining », Thèse de doctorat, Université de Québec à Trois-Rivières, Canada, 2007.
- [Joseph, 2014]** F.Joseph, « Enjeux et défis de la contribution des réseaux sociaux numériques à une transmission réussie: le cas de l'Eglise catholique », thèse de doctorat, Université de Reims Champagne-Ardenne, Reims, France, 2014.
- [Kacem et al., 2014]** A.Kacem, M.Boughanem, R.Faiz, « Time-Sensitive User Profile for Optimizing Search Personalization », In V. Dimitrova, T. Kuflik, D. Chin, F. Ricci, P. Dolog, et G.-J. Enrichissement du profil utilisateur à partir de son réseau social 16 Houben (Eds.), User Modeling, Adaptation, and Personalization, p. 111!121. Springer International Publishing, 2014.
- [Kaplan et al., 2010]** A.Kalpan, M.Haenlein, « Users of the world, unite! The challenges and opportunities of social media », Journal Business horizons, 2010.
- [Kenter et al., 2018]** T.Kenter, A.Borisov, V.G.Christophe, M.Dehghani, R.Maarten, M.Bhaskar, « Neural Networks for Information Retrieval », In WSDM 2018: The Eleventh ACM International Conference on Web Search and Data Mining, CA, USA, ACM, New York, USA, 2018.

- [Kim et al., 2003]** H.R.Kim, P.K.Chan, « Learning implicit user interest hierarchy for context in personalization », In Proceedings of the 8th International Conference on Intelligent User Interfaces, IUI 2003, New York, USA, 2003.
- [Kim et al., 2011]** H.N.Kim, A.Elkhaldi, A. El Seddik, G.S.Jo, « Collaborative user modeling with user-generated tags for social recommender systems », Expert Systems with Applications, vol. 38, 2011.
- [King et al., 2016]** J.King, J.Forder, « Data analytics and consumer profiling: finding appropriate privacy principals for discovered data », in: Computer Law and Security Review-CLSR, vol.32, 2016.
- [Kuflik et al. 2006]** T.Tuflik, Z.Boger, P.Shoval, « Filtering search results using an optimal set of terms identified by an artificial neural network », an International Journal: Information Processing and Management 42(2), 2006.
- [Kumar et al., 2008]** R.Kumar, L.Backstrom, J.Leskovic, A.Tomkins, « Microscopic evolution of social networks », In KDD'8: Proceeding of the 14 ACM SIGKDD international conference on knowledge discovery and data mining, New York, USA, 2008.
- [Li et al., 2013]** D.Li, P.Cao, Y.Guo, M.Lei, « Time Weight Update Model Based on the Memory Principle in Collaborative Filtering », Journal of Computers 8, 2013.
- [Lops et al., 2011]** P.Lops, M.D.Gemmis, G.Semeraro, « Content-based Recommender Systems: State of the Art and Trends », RICCI F., ROKACH L., SHAPIRA B., KANTOR P.B. (eds.), Recommender Systems Handbook, Springer US, 2011.
- [Ma et al., 2011]** Y.Ma, Y.Zeng, X.Ren, N.Zhong, « User interests modeling based on multi-source personal information fusion and semantic reasoning », In Proceedings of the 7th international conference on active media technology, Chine, 2011.
- [Maloof et al., 2000]** M.A.Maloof, R.S.Michalski, « Selecting Examples for Partial Memory Learning », Machine Learning, 41, In Proceedings of the Seventh IEEE International Conference on Tools with Artificial Intelligence, Netherlands, Pays-Bas, 2000.
- [Massa et al., 2007]** P.Massa, P.Avesani, « Trust-aware Recommender Systems », In Proceedings of the 2007

- ACM Conference on Recommender Systems, New York, NY, USA, 2007.
- [Meier, 2011]** P.Meier, « Protection des données- Fondements », Edition principes généraux et droit privé, 2011.
- [Meo et al., 2010]** P.D.Meo, G.Quattrone, D.Ursino, « A query expansion and user profile enrichment approach to improve the performance of recommender systems operating on a folksonomy », User Modeling and User-Adapted Interaction, vol. 20, 2010.
- [Mercier, 2008]** P.A.Mercier , « Liens faibles sur les courants faibles », International Journal informations sociales, 2008.
- [Moukas, 1997]** A.Moukas, « Amalthea: Information discovery and filtering using a multiagent evolving ecosystem », An International Journal: Applied Artificial Intelligence, 1997.
- [On-At, 2017]** S.On-At, «Temporalité et réseaux sociaux : prise en compte de l'évolution dans la construction du profil utilisateur », Thèse de doctorat, Université de Toulouse, France, 2017.
- [Ramiandrisoa et al., 2017]** F.Ramiandrisa, J.Mothe, « Profil utilisateur dans les réseaux sociaux : État de l'art », International conférence Rencontres Jeunes Chercheurs en Recherche d'Information, Marseille, France, 2017.
- [Renders, 1995]** J.M. Renders, « Algorithmes génétiques et Réseaux de Neurones », Editions HERMES, 1995.
- [Ricci et al., 2011]** F.Ricci, L.Rokach, B.Shapira, P.B.Kantor, « Recommender Systems Handbook », Edion computer science, 1, Springer, 2011.
- [Roth et al., 2010]** M.Roth, A.Ben-David, D.Deutsher, G.Flysher, I.Horn, A.Leichtberg, « Suggesting friends using the implicit social graph », In Proceedings of the 16th acm sigkdd international conference on knowledge discovery and data mining, New York, USA, 2010.
- [Sargsyan, 2017]** M.Sargsyan, « Endorsements and Behavioral Advertising in Social Media under EU, Swiss and US law - Disclosure requirements, personality rights, and data protection », Center for Information Technology Society and Law (ITSL), vol.3, 2017.

- [Spiliopoulou, 2011]** M. Spiliopoulou, « Evolution in Social Networks: A Survey », In C. C. Aggarwal (Ed.), Social Network Data Analytics, 2011.
- [Sueur et al., 2011]** C.Sueur, J.L.Deneubourg, « Self-Organization in Primates: Understanding the Rule Underlying Collective Movements », International Journal of Primatology, 2011.
- [Tang et al., 2010]** L.Tang, H.Lui, « Community detection and mining in social media », Morgan and Claypool Publishers, 2010.
- [Tchunte et al., 2013]** D.Tchunte, F.Canut, N.Jessel, A.Paninou, F.Sades, « A community-based algorithm for deriving users profiles from egocentric networks: experiment on Facebook and DBLP », Social Network Analysis and Mining, 2013.
- [Tuffery, 2012]** S. Tuffery, « Data Mining et statistique décisionnelle », Editions TECHNIP, 2012.
- [Vallet et al., 2010]** D.Vallet, I.Cantador, J.M.Jose, « Personalizing web search with folksonomy-based user and document profiles », In C. Gurrin et al. (Eds.), Advances in information retrieval, vol 5993, 2010.
- [Walter, 2011]** J.P.Walter, « Le profilage des individus à l'heure du cyber espace - un défi pour le respect du droit à la protection des données », Edition L'interconnexion de données : Problématique et cadre juridique, 87, 2011.
- [Wang et al., 2010]** J.Wang, M.Clements, J.Yang, A.P.Vries, M.J.T.Reinders, « Personalization of tagging systems », Inf. Process. Manage vol. 46, 2010.
- [White et al., 2009]** R.W.White, P.Bailey, L.Chen, « Predicting user interests from contextual information », In Proceedings of the 32Nd international ACM SIGIR conference on research and development in information retrieval, New York, USA, 2009.
- [Zemirli et al., 2005]** W.N.Zemirli, L.Tamine, M.Boughanem, « Accès personnalisé à l'information : vers la définition d'un profil utilisateur multidimensionnel », In international Symposium On Programming Systems (ISPS, 2005), USTHB, Alger, Algérie, 2005.
- [Zemirli, 2008]** W.N.Zemirli, « Modèle d'accès personnalisé à l'information basé sur les Diagrammes d'influence intégrant un profil utilisateur évolutif », Thèse de doctorat, Université Paul Sabatier de Toulouse 3, France, 2008.

- [Zhang et al., 2010]** B.Zheng, Y.Guan, H.Sun, Q.Liu, J.Kong, « Survey of user behaviors as implicit feedback », In 2010 international conference on computer, mechatronics, control and electronic engineering (CMCE), vol. 6, Chine, 2010.
- [Zheng et al., 2011]** N.Zheng, Q.Liu, « A recommender system based on tag and time information for social tagging systems », Expert Syst. Appl., vol. 38, no 4, 2011.